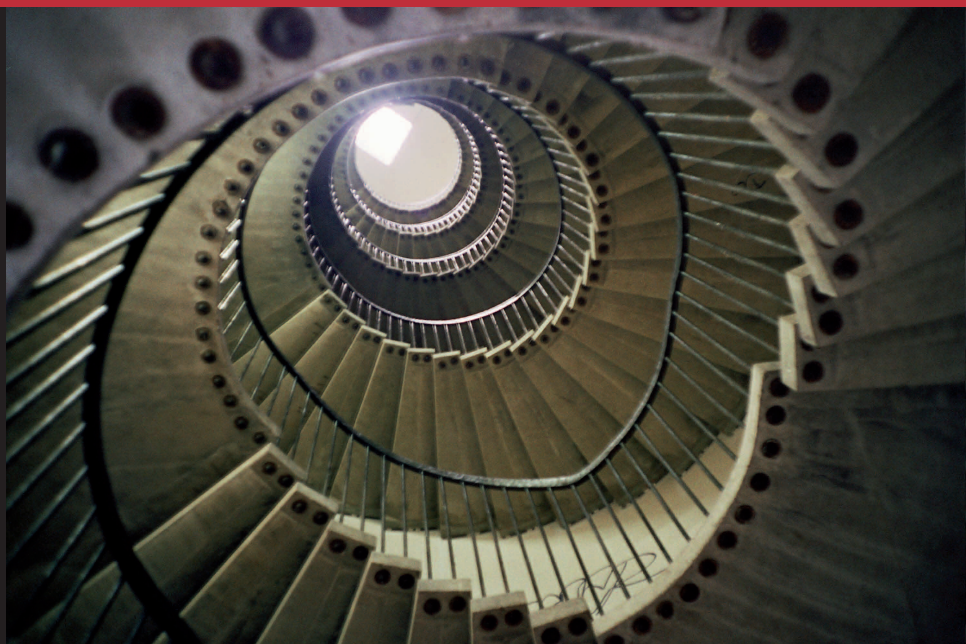


Autour des langues et du langage



Perspective pluridisciplinaire

Coordonné par Mathieu Loiseau, Myriam Abouzaïd, Laurence Buson, Cristelle Cavalla, Ali Djaroun, Céline Dugua, Anna Ghimenton, Vannina Goossens, Thomas Lebarbé, Aurélie Nardy, Fanny Rinck, Christian Surcouf



**Autour des langues et du langage :
Perspective pluridisciplinaire**

Photographies de couverture : Vannina Goossens

ISBN 978-2-7061-1427-4



Le code de la propriété intellectuelle n'autorisant, aux termes de l'article L.122-5, 2° et 3° a, d'une part, que les « copies ou reproductions strictement réservées à l'usage privé du copiste et non destinées à une utilisation collective » et, d'autre part, que les analyses et les courtes citations dans un but d'exemple et d'illustration, « toute représentation ou reproduction intégrale ou partielle faite sans le consentement de l'auteur ou de ses ayants droit ou ayants cause est illicite » (art. L.122-4).

Cette représentation ou reproduction, par quelque procédé que ce soit, constituerait donc une contrefaçon sanctionnée par les articles L.335-2 et suivants du code de la propriété intellectuelle.

AUTOUR DES LANGUES ET DU LANGAGE : PERSPECTIVE PLURIDISCIPLINAIRE

Ouvrage coordonné par :

Mathieu LOISEAU, Myriam ABOUZAÏD, Laurence BUSON,
Cristelle CAVALLA, Ali DJAROUN, Céline DUGUA,
Anna GHIMENTON, Vannina GOOSSENS, Thomas LEBARBÉ,
Aurélié NARDY, Fanny RINCK, Christian SURCOUF

AVANT-PROPOS

par Jacqueline Billiez

Du 4 au 7 juillet 2006 s'est tenu à l'université Stendhal-Grenoble 3 un *colloque international d'étudiants chercheurs en didactique des langues et linguistique*, organisé par une dizaine de doctorants guidés par deux jeunes maîtres de conférences, tous membres du laboratoire LIDILEM (LInguistique et DIdictique des Langues Etrangères et Maternelles), auquel se rattachaient à ce moment-là environ 80 doctorants et une quarantaine d'enseignants-chercheurs de cette université, de l'IUFM de Grenoble et d'autres IUFM du quart sud-est hexagonal.

Dès l'idée lancée et avalisée par une réunion du laboratoire, un peu plus d'un an avant, le comité d'organisation qui s'était constitué a pris en mains, selon un rythme d'une intensité extrême, toutes les opérations nécessaires dans l'organisation d'un colloque de cette envergure : constitution d'un comité scientifique ; montage des dossiers pour l'obtention de labels et de subventions ; contacts avec des conférenciers potentiels ; organisation du logement, de la restauration, etc. ; rédaction d'un appel à communication qui insistait sur le caractère très pluridisciplinaire des thématiques possibles pour des propositions de communication et mettait aussi l'accent sur le caractère international (avec usage possible de l'anglais) de la rencontre d'une part, et, d'autre part, sur l'aspect « formation à la proposition de communication puis à la rédaction d'articles » des apprentis chercheurs par des enseignants-chercheurs et chercheurs relecteurs experts dans les divers champs couverts par les domaines linguistique et didactique ; création d'un site Web très élaboré destiné à la soumission des communications et à l'interface entre experts/apprentis/comité d'organisation.

Le dispositif mis en place a généré une motivation qui a largement dépassé les prévisions en termes d'accueil de communicants. En effet, malgré un calendrier serré (et respecté à la lettre), une procédure de soumission lourde et exigeante, le comité d'organisation a reçu 254 propositions de communications, en provenance de jeunes chercheurs de 72 équipes de recherche hébergées dans 32 universités différentes et implantées sur les 5 continents. L'organisation de la sélection des soumissions était donc impérative et, je dois dire qu'elle a été gérée avec une rigueur et une efficacité que je n'avais jamais rencontrée auparavant lors des sélections de communications auxquelles j'avais pu participer. Plus de 100 enseignants-chercheurs ont accepté de participer aux relectures des soumissions et chaque auteur a bénéficié de deux relecteurs, avec des conseils de chacun pour la suite fût-elle heureuse ou malheureuse.

Au total ce sont environ 100 communications orales qui ont été proposées dans 3 sessions en parallèle, regroupées selon des thématiques très diversifiées comme : interactions en classe, neurolinguistique et pathologies, passage L1/L2, lexicologie et terminologie, développement syntaxique, sociolinguistique et multilinguisme, sociolinguistique et variation, lexique et développement, analyse du discours, sociolinguistique et représentations, didactique de l'écrit, outils pour la didactique, interlangue et interférence, syntaxe, sociolinguistique et ethnologie, traduction, sociolinguistique et didactique, pratiques des apprenants, écriture/lecture, écrit et nouvelles technologies, le verbe, traitement automatique des langues. Sans oublier une trentaine de posters qui ont été exposés et surtout présentés en fin de session en 2 minutes par chaque auteur dans le but d'inciter à la visite. Objectif qui a été effectivement atteint. Les auteurs ne sont pas restés inactifs devant leur poster : ils ont été sollicités au même titre que les communicants en sessions, qui, eux-mêmes – fait rarissime – n'ont jamais dépassé leur temps de communication/discussion imparti grâce à une double animation de session par un enseignant-chercheur du Lidilem assurant l'animation scientifique et un étudiant-chercheur gardien du temps sans baisse de vigilance ni laxisme dans le respect des consignes, ce qui a rendu possibles et non-gênantes les migrations d'un amphi à l'autre.

Les membres du comité d'organisation ont tout pensé et organisé dans les moindres détails à partir de tout un travail de réflexion sur les défauts et dysfonctionnements divers bien connus des colloques de seniors auxquels ils avaient déjà participé. Ils ont en conséquence prévu toutes les parades pour les éviter et améliorer ainsi la qualité scientifique attendue de ce genre d'événement.

Chaque journée a en outre débuté par une conférence suivie d'un temps de discussion :

François RASTIER a ainsi ouvert le colloque sur Langues et corpus - de nouveaux observables et de nouveaux objectifs pour la linguistique ?, Harriet JISA a traité du développement tardif : devenir plus qu'un locuteur natif, Liam MURRAY a été invité à parler *In search of a personalisation of evaluation in CALL contexts* et Françoise GADET du style dans la dynamique des langues.

Le buffet de clôture a lui aussi laissé des souvenirs inoubliables tant pour la qualité des mets que pour celle de la participation des membres du comité d'organisation, qui a été chaleureusement félicité à la fois par les étudiants-chercheurs communicants, participants et les membres du Lidilem. La directrice que j'étais à l'époque les a remerciés pour ce beau cadeau qui lui était en quelque sorte offert à la veille de son passage de relais pour assurer la direction du Lidilem.

Le dernier travail – et non des moindres – de ce comité a concerné la publication de ce recueil. La centaine d'enseignants-chercheurs, mise à contribution pour la sélection des communications et posters, a effectué une nouvelle relecture des articles reçus avant la date limite selon le même protocole que pour les soumissions. Au final sur les 92 articles reçus, 67 ont été retenus pour cet ouvrage, organisé en 5 chapitres aux thématiques larges qui, dans l'esprit du Lidilem, n'excluent pas les recouvrements : les descriptions linguistiques, la psycholinguistique et le développement langagier, la sociolinguistique et le plurilinguisme, la didactique des langues, les outils informatiques pour la linguistique et la didactique des langues.

Seul un titre ouvert sur la diversité des objets, des disciplines et des approches de recherche pouvait rendre compte de la richesse et de la qualité des contributions sélectionnées.

Autour des langues et du langage : perspectives pluridisciplinaires est un ouvrage qui fera date par la dynamique qui s'est révélée à l'occasion de ce colloque dont tous les participants enthousiastes se sont séparés en étant persuadés qu'il ne restera pas sans suite.

Apprendre à communiquer sur sa recherche et, étape supérieure, à organiser un colloque font pleinement partie de la formation des étudiants-chercheurs. Les accompagner représente alors une mission des enseignants-chercheurs et chercheurs confirmés, et les soutenir, celle d'instances comme les équipes de recherche, les écoles doctorales et les centres d'initiation à l'enseignement supérieur (CIES).

REMERCIEMENTS

Les coordinateurs de cet ouvrage tiennent à remercier les nombreux relecteurs sans lesquels ce recueil n'aurait pas vu le jour :

Comité scientifique :

BILLIEZ Jacqueline, Université Stendhal Grenoble III (présidente)
BINON Jean, Université Catholique de Louvain, Belgique
BRONCKART Jean-Paul, Université de Genève, Suisse
DE CHANAY Hugues, Université Lumière Lyon II
CHANIER Thierry, Université de Franche-Comté
CHISS Jean-Louis, Université de la Sorbonne Nouvelle Paris 3
GADET Françoise, Université Paris 10 Nanterre
JISA Harriet, Université Lumière Lyon II
MURRAY Liam, University of Limerick, Irlande
POLGUÈRE Alain, Université de Montréal, Canada
RASTIER François, CNRS, Paris

Comité de lecture :

ABOUDA Lotfi, Université d'Orléans
ABRY Dominique, CUEF, Grenoble3
AIT-MOKHTAR Salah, Xerox XRCE, Grenoble
ANTONIADIS Georges, Lidilem, Grenoble 3
ARROYO-GONZALEZ Encarnacion, Laboratoire Jacques Lordat, Toulouse 2
BARBU Stéphanie, CNRS, Université Rennes 1
BARRÉ DE MINAC Christine, IUFM, Lidilem, Grenoble 3
BENRABAH Mohamed, Université Stendhal, Grenoble3
BIANCO Maryse, LSE, Université Grenoble 2
BILLIEZ Jacqueline, Lidilem, Grenoble 3
BINON Jean, Université Catholique de Louvain
BLANC Nathalie, IUFM, Lyon
BOCH Françoise, Lidilem, Grenoble 3
BOHAS Georges, responsable de la section Langues sémitiques à l'ENSLSH
BOTÉVA Sylvia, Université de Sofia
BOUCHARD Robert, ICAR, Lyon 2
BOURQUIN Céline, Université de Lausanne

BRISAUD Catherine, IUFM, LIDILEM, Grenoble3
BRODIN Elisabeth, LIDILEM
BRONCKART Jean-Paul, Université de Genève
CALAQUE Elisabeth, Lidilem, Grenoble 3
CANELAS-TREVISI Sandra, Lidilem, Grenoble 3
CARPITELLI Elisabetta, Centre de dialectologie - U-Grenoble 3
CARRAS Catherine, LIDILEM
CAVALLA Cristelle, Lidilem, Grenoble 3
CHANIER Thierry, Université de Franche-Comté
CHAUVIN Carole, Centre de Dialectologie de Grenoble, Université Stendhal
CHENU Florence, CNRS, Lyon 2
CHEVROT Jean-Pierre, Lidilem, Grenoble 3
CHISS Jean-Louis, Université de la Sorbonne Nouvelle, Paris 3
COLLETTA Jean-Marc, Lidilem, Grenoble 3
CREISSELS Denis, DDL, Lyon 2
CUQ Jean-Pierre, Département Sciences du Langage, Sofia Antipolis
DABÈNE Michel, Lidilem, Grenoble 3
DABÈNE Louise, Lidilem, Grenoble 3
DAUNAY Bertrand, Université Lille 3
DE CHANAY Hugues, Université Lumière, Lyon 2
DEJEAN Charlotte, LIUM, Le Mans
DEVELOTTÉ Christine, ICAR, Lyon 2
DOMPMARTIN-NORMAND Chantal, Lidilem, Grenoble 3
DUVIGNAU Karine, Laboratoire Lordat, U. Toulouse 2 & IUFM Midi Pyrénées
FAGYAL Zsuzsanna, University of Illinois at Urbana-Champaign
FRIER Catherine, Lidilem, Grenoble 3
GADET Françoise, Université Paris 10, Nanterre
GAJO Laurent, ELCF, Genève
GALLIGANI Stéphanie, Lidilem, Grenoble 3 - Diltec, Paris 3
GASQUET-CYRUS Médéric, Université de Provence - Aix-Marseille I
GROSSMANN Francis, LIDILEM, Grenoble
GUERNIER Marie-Cécile, Lidilem, Grenoble 3
GUIDETTI Michèle, Laboratoire Cognition, Communication et Cognition, Communication et Développement
HAMBYE Philippe, UCL
HENDERSON Alice, Université de Savoie
HUDELOT Christian, CNRS - Université Paris 5
JACQUET Guillaume, XRCE
JEANNERET Thérèse, ILCF, Lausanne
JISA Harriet, DDL, Lyon 2
KRAIF Olivier, Lidilem, Grenoble 3
LACHERET-DUJOUR Anne, Laboratoire CRISCO, Université de Caen
LAGORGETTE Dominique, Laboratoire "Littératures Langues Sociétés", Université de Savoie
LEBARBÉ Thomas, Lidilem, Grenoble 3
LEDEGEN Gudrun, LCF-UMR 8143 du CNRS, Université de la Réunion
LEPOIRE-DUC Solveig, ICAR, Lyon 2
MANGENOT François, Lidilem, Grenoble 3
MANGIANTE Jean-Marc, Chambre de Commerce et d'Industrie de Paris
MARSICO Egidio, DDL, Lyon 2
MARTINIE Bruno, DDL, Lyon 2
MASPERI Monica, Lidilem, Grenoble 3

MATTHEY Marinette, Lidilem, Grenoble 3
 MAZAURIC Catherine, LLA, Toulouse 2
 MILLET Agnès, Lidilem, Grenoble 3
 MOÏSE Claudine, Université d'Avignon
 MOORE Danièle, Simon Fraser University
 MURRAY Liam, University of Limerick, Limerick, Irlande
 NOVAKOVA Iva, Lidilem, Grenoble 3
 PACTON Sébastien, Université Paris 5 et CNRS
 PANIER Louis, ICAR, Lyon 2
 PARPETTE Chantal, ICAR, Lyon 2
 PELLENQ Catherine, IUFM, LSE Grenoble 2
 PICCARDO Enrica, CEDICLEC, St Etienne - IUFM, Grenoble
 PIEROZAK Isabelle, Université de Picardie Jules Verne / U.P.J.V.
 POLGUÈRE Alain, Département de linguistique et traduction, Montréal
 POLLET Marie-Christine, Université libre de Bruxelles
 PONTON Claude, Lidilem, Grenoble 3
 PRUDENT Lambert-Félix, Université de la Réunion
 RASTIER François, CNRS, Paris
 RÉMI Sylvianne, ICAR, Lyon 2
 RISPAIL Marielle, IUFM, Lidilem, Grenoble 3
 SAUTOT Jean-Pierre, Lidilem, Grenoble 3
 SCHELSTRAETE Marie-Anne, Université Catholique de Louvain
 SEGOND Frédérique, Xerox XRCE, Grenoble
 SIMON Jean-Pascal, IUFM, Lidilem, Grenoble 3
 SIMON Diana-Lee, IUFM, Lidilem, Grenoble 3
 SOUBRIÉ Thierry, Lidilem, Grenoble 3
 SPINELLI Elsa, LPNC, Grenoble 2
 TOTEREAU Corinne, IUFM, LIDILEM, Grenoble3
 TRIMAILLE Cyril, Lidilem, Grenoble 3
 TUTIN Agnès, Lidilem, Grenoble 3
 VALLÉE Nathalie, ICP, Grenoble
 VENEZIANO Edy, Université Paris 5 - CNRS
 VERLINDE Serge, Université Catholique de Louvain, Belgique
 WILLIAMS Geoffrey, Université Bretagne Sud, Lorient

Sans oublier les membres du comité d'organisation du colloque :

Myriam ABOUZAÏD, Luc BIICHLÉ, Alexia BLANCHARD, Katia BOUCHOUÉVA,
 Laurence BUSON, Cristelle CAVALLA, Ali DJAROUN, Céline DUGUA, Anna GHIMENTON,
 Vannina GOOSSENS, Patricia LAMBERT, Thomas LEBARBÉ, Mathieu LOISEAU,
 Nathaël MARTIN, Saskia MUGNIER, Aurélie NARDY, Coralie PAYRE-FICOUT,
 Karine PERRIER-ROSSET, Luu Dinh PHUC, Severine RAPANOËL, Fanny RINCK,
 Christian SURCOUF, Nathalie THAMIN & Thi Thu Ha TRAN

TABLE DES MATIÈRES

DESCRIPTION LINGUISTIQUE

VARIABLES DE PERSONNALITÉS ET USAGE DU LEXIQUE ÉMOTIONNEL PAR DES ÉTUDIANTS SCIENTIFIQUES ET LITTÉRAIRES <i>Rachid BANNOUR</i>	1
UNE DIFFICULTÉ DE LA DESCRIPTION LINGUISTIQUE : L'OBJET DIRECT EN FRANÇAIS <i>Florinela COMANESCU</i>	9
STRATÉGIES DE RESTITUTION DES CONSTRUCTIONS VERBALES SÉRIELLES DU CRÉOLE HAÏTIEN EN FRANÇAIS L2 <i>Darline COTHIÈRE-ROBERT</i>	17
LES PROCÉDÉS DE MÉTAPHORISATION DANS LE DISCOURS MÉDICAL ARABE : ÉTUDE DE CAS <i>Tatiana EL-KHOURY</i>	25
LE CONCEPT DE PHRASE CANONIQUE EN LINGUISTIQUE FRANÇAISE <i>Antoine GAUTIER</i>	33
SUR LA VALEUR "PROTOTYPANTE" DU PRÉSENT DE L'INDICATIF <i>Rose-Marie GERBE</i>	39
ANALYSE TEXTUELLE, ANALYSE DE GENRES : QUELLES RELATIONS, QUELS INSTRUMENTS ? <i>Mathilde GONÇALVES, Florencia MIRANDA</i>	47
RÉGULARITÉS ET IRRÉGULARITÉS DANS LA POLYSÉMIE DES NOMS D'AFFECT : LE CAS DE L'EXPRESSION DE LA SOURCE DE L'AFFECT <i>Vannina GOOSSENS</i>	55
LA THÉORIE DES MATRICES ET DES ÉTYMONS EN BERBÈRE OU NOUVELLE HIÉRARCHISATION DU LEXIQUE <i>Cécile LUX</i>	63
POUR UNE ÉTUDE DE L'APOSTROPHE INJURIEUSE <i>Iuliana-Anca MATEIU</i>	71
LA DIMENSION INDIVIDUELLE DE LA PAROLE : APPROCHES LINGUISTIQUE ET COGNITIVE <i>Séverine MORANGE</i>	79
THE PRODUCTION OF CORONAL CONSONANTS BY FRENCH LEARNERS OF ENGLISH <i>Stéphane MORTREUX</i>	87

ESSAI DE DESCRIPTION D'EMPLOIS DES MOTS DU DISCOURS AUTOUR DE <i>HEIN</i> <i>Hiroko NODA</i>	95
LAZY AGREEMENT IN THE LADIN DP <i>Sabrina RASOM</i>	103
LES CRITÈRES DE GRAMMATICALISATION POUR LES TEMPS COMPOSÉS EN FRANÇAIS ET EN ALLEMAND <i>Aude REBOTIER</i>	111
ÉTUDE TYPOLOGIQUE DE L'EXPRESSION DU DÉPLACEMENT EN ARABE TUNISIEN <i>Darine SAÏDI</i>	119
LA MOTIVATION DANS LES NOMS DES MALADIES EN ROUMAIN : DES EXEMPLES DE TABOU LINGUISTIQUE <i>Carmen SCARLAT</i>	127
PHONETIC ALIGNMENT AND THE PHRASE TONE / 'TRAILING' TONE AMBIGUITY IN IRISH ENGLISH <i>Jennifer SULLIVAN</i>	135
LE PRÉSENT EXPRIME-T-IL L'ACTUALITÉ ? LA DOUBLE PERSPECTIVE DE L'ENCODAGE ET DU DÉCODAGE <i>Christian SURCOUF</i>	143
LA VARIATION DIATOPIQUE DU FRANÇAIS DANS L'ŒUVRE LITTÉRAIRE D'YVES VIOLLIER (VENDÉE) <i>Inka WISSNER</i>	151

PSYCHOLINGUISTIQUE ET DÉVELOPPEMENT LANGAGIER

TERMES DE COULEUR EN DISCOURS ET CONCEPTS DE COULEUR <i>Caroline CANCE</i>	161
ÉVALUATION DE LA REPRÉSENTATION INTRA- ET INTER-HÉMISPHERIQUE DU LANGAGE CHEZ LES SUJETS SAINS : ÉTUDE COMPORTEMENTALE EN CHAMP VISUEL DIVISÉ ET ÉTUDE IRMF <i>Émilie COUSIN</i>	169
ÉTUDE DU LANGAGE SPONTANÉ D'ENFANTS SOURDS MOYENS ET LÉGERS : VERS UN ÉVITEMENT DE LA COMPLEXITÉ SYNTAXIQUE? <i>Hélène DELAGE</i>	177
LES TROUBLES PRAGMATIQUES ET ÉMOTIONNELS CHEZ DES ENFANTS ET ADOLESCENTS CEREBROLESES : L'EXEMPLE DE LA SITUATION DE PRODUCTION NARRATIVE <i>Anaïg DELANOË</i>	185
FORMES LINGUISTIQUES ADJECTIVALES ET IDENTIFICATION DES REPRÉSENTATIONS COGNITIVES <i>Gaëlle DELEPAUT</i>	193
"UN OURS" / "DES NOURS" OU LE RÔLE DE LA FRÉQUENCE SUR L'ACQUISITION DE LA LIAISON EN FRANÇAIS <i>Céline DUGUA</i>	201
MODALISATION ET APPROXIMATIONS SÉMANTIQUES ENTRE 2 ET 8 ANS : POUR UNE DICHOTOMIE NOMS VS VERBES <i>Juliette ÉLIE</i>	209
LE DÉVELOPPEMENT MÉTAPHONOLOGIQUE CHEZ LES ENFANTS BILINGUES ET MONOLINGUES <i>Angélique LAURENT</i>	217

LE MARQUAGE ASPECTUEL DE LA SIMULTANÉITÉ CHEZ DES APPRENANTS QUASI-NATIFS FRANCOPHONES DE L'ANGLAIS DANS UNE TÂCHE NARRATIVE	
<i>Pascale LECLERCQ</i>	225
LE RÔLE DE LA CONVERSATION DANS LA TRANSITION DES ÉNONCÉS À UN MOT VERS LES ÉNONCÉS À DEUX MOTS : UNE ÉTUDE EXPÉRIMENTALE	
<i>Silvia NIEVA</i>	233
SYNTAXE, ACQUISITION ET DIALECTOLOGIE : LA PLURIDISCIPLINARITÉ APPLIQUÉE À L'ÉTUDE DU SUJET NUL ENFANTIN	
<i>Katérina PALASIS-JOURDAN</i>	241
FORMES PAR DÉFAUT ET FRÉQUENCE DANS L'INPUT EN FLE	
<i>Anita THOMAS</i>	249

SOCIOLINGUISTIQUE ET PLURILINGUISME

LA STANDARDISATION DE L'AMAZIGHE (BERBÈRE) AU MAROC : ENJEUX ET RÉCEPTION AUPRES D'ENSEIGNANTS	
<i>Myriam ABOUZAÏD</i>	259
REPRÉSENTATIONS DE MIGRANTS ORIGINAIRES DU MAGHREB : DISTINCTION DES VARIÉTÉS DIGLOSSIQUES PAR LES REPRÉSENTATIONS SUR LA PRODUCTION ET LA RÉCEPTION	
<i>Luc BIICHLÉ</i>	267
INFLUENCE DU MILIEU SOCIAL ET DU RÉSEAU DE PAIRS SUR LES REPRÉSENTATIONS DE LA VARIATION STYLISTIQUE CHEZ DES ENFANTS PRÉADOLESCENTS FRANÇAIS	
<i>Laurence BUSON</i>	275
L'ENVIRONNEMENT LANGAGIER DU TRÈS JEUNE ENFANT DANS LA SOCIÉTÉ PLURILINGUE DE VÉNÉTIE : DESCRIPTION D'INDICES STATISTIQUES ET PRAGMATIQUES	
<i>Anna GHIMENTON</i>	283
LES DÉTERMINANTS DANS L'INTERLANGUE DES BILINGUES TURC-FRANÇAIS ET TURC- ANGLAIS	
<i>Jeanne GONAC'H</i>	291
USAGES LANGAGIERS ENFANTINS ET VARIABLES SOCIOLINGUISTIQUES : CONVERGENCES MICROSOCIOLOGIQUES ET STRATIFICATION MACROSOCIOLOGIQUE	
<i>Aurélie NARDY, Nathaël MARTIN</i>	299
LES PYGMÉES BAKA DU GABON : APPROCHE SOCIOLINGUISTIQUE	
<i>Pascale PAULIN</i>	307
CONTACTS DES LANGUES ET IDENTITÉ : LE RÔLE DE L'INNOVATION LEXICALE DANS LE CAS DU SCAMTO	
<i>Lerato Caroline POOE</i>	315
DES REPRÉSENTATIONS DIGLOSSIQUES AUX PRATIQUES INTERLECTALES : CONFRONTATION DE CONCEPTS SOCIOLINGUISTIQUES EN CONTEXTE RÉUNIONNAIS	
<i>Séverine RAPANOËL</i>	323
COMPÉTENCE PLURILINGUE, REPRÉSENTATIONS DES LANGUES ET CONSTRUCTION DES CONNAISSANCES CHEZ LES APPRENANTS FRANCO-ALLEMANDS	
<i>Sofia STRATILAKI</i>	331
LA PLACE ET LES FONCTIONS RÉSERVÉES À LA LANGUE MINORÉE : LE CAS DE BAFRA (GRECE)	
<i>Maria ZERVA</i>	339

DIDACTIQUE DES LANGUES

CARACTÉRISATION DE LA NOTION D'ERREUR LEXICALE ET DESCRIPTION DES PROBLÈMES LEXICAUX D'ÉTUDIANTS UNIVERSITAIRES

Dominic ANCTIL349

METTRE EN ŒUVRE SES COMPÉTENCES DISCURSIVES EN L2 : QUELLES OPPORTUNITÉS EN MILIEU SCOLAIRE ?

Évelyne BERGER.....357

LES PROCÉDÉS D'AUTOCOPIE DE L'ENFANT APPRENTI SCRIPTEUR

Jocelyne CUSSAC POMEL, Geneviève GARCIA BOURGADE365

COMMENT DIRE OUI ? LES MARQUES D'ACCORD DANS LE CORPUS LANCOM ANALYSE DIDACTIQUE ET LINGUISTIQUE

Juliette DELAHAIE373

QUAND LES LYCÉENS ALGÉRIENS PARLENT DE LEURS COMPÉTENCES SCRIPTURALES : ENJEUX ET PERSPECTIVES DIDACTIQUES

Ali DJAROUN.....381

ÉTUDE ET OPÉRATIONNALISATION DE PHÉNOMÈNES DE STÉRÉOTYPIE POUR LA DIDACTIQUE DE L'ÉCRITURE

Bernadette KERVYN.....389

OUTILS INFORMATIQUES POUR LA LINGUISTIQUE ET LA DIDACTIQUE DES LANGUES

ESQUISSE D'UNE FORMALISATION DES ACTES DE LANGAGE INDIRECTS

Sophie ANQUETIL399

CONSTITUTION DE RESSOURCES PÉDAGOGIQUES NUMÉRIQUES : LE LEXIQUE DES AFFECTS

Magdalena AUGUSTYN, Sabrina BEN HAMOU, Gwendoline BLOQUET, Vannina GOOSSENS, Mathieu LOISEAU, Fanny RINCK.....407

CONTRIBUTION DES CORPUS À L'ENRICHISSEMENT DES DICTIONNAIRES BILINGUES GÉNÉRAUX. APPLICATION AU DOMAINE DE LA VOLCANOLOGIE.

Cécile FRÉROT, Amélie JOSSELIN-LERAY.....415

ORGANISATION DES FONCTIONS LEXICALES POUR UN MEILLEUR ACCÈS À L'INFORMATION DANS LE DiCo

Anne-Laure JOUSSE423

APPROCHE COMPARÉE DES INTERACTIONS ENTRE APPRENANTS DANS LA CLASSE DE LANGUE ET DANS LES ENVIRONNEMENTS PÉDAGOGIQUES INFORMATISÉS

Barbara RAZOLA MAYOR.....431

CALCUL DYNAMIQUE DU SENS D'UN ADJECTIF : DE L'ANTE À LA POSTPOSITION, UN MÉCHANT CHANGEMENT DE SENS.

Fabienne VENANT439

LA PLATE-FORME LINGUAStream

Antoine WIDLÖCHER, Frédéric BILHAUT.....447

DESCRIPTION LINGUISTIQUE

VARIABLES DE PERSONNALITÉS ET USAGE DU LEXIQUE ÉMOTIONNEL PAR DES ÉTUDIANTS SCIENTIFIQUES ET LITTÉRAIRES

Rachid BANNOUR

rachid.bannour@up.univ-aix.fr

*Centre de Recherche en Psychologie de la Connaissance, du Langage et de l'Emotion
Université de Provence, 29, avenue Robert Schuman,
13621 Aix-en-Provence cedex 1*

Abstract : This study examines the effect of several personality variables on writing activities among university students. When writing on emotional topics and neutral ones, students' use of emotional lexis differed in relation to their broader field of study. The writings of arts and literature students revealed two orientations, related to perfectionism and to procrastination, whereas the writings of science students were linked to anxiety and apprehension.

1. Cadrage théorique

La psychologie cognitive s'emploie à repérer les différents processus et connaissances qui entrent en jeu dans l'activité rédactionnelle (Hayes, 1996 ; Olive & Piolat, 2003). Seules, quelques rares études concernant la maîtrise de la production écrite prennent en compte les difficultés qu'une partie de la population estudiantine éprouve. Ces difficultés peuvent aboutir à une appréhension à écrire, parfois même un blocage, au point d'éviter certains cursus académiques. Selon Boice (1985), ces difficultés auraient pour origine des causes de nature cognitivo-émotionnelle comme l'appréhension à écrire, la procrastination, le perfectionnisme, la dysphorie, l'impatience, la peur de l'évaluation et le contrôle des règles.

Dans le cadre de la présente recherche, seuls trois de ces facteurs seront pris en compte: l'appréhension à écrire, la procrastination et le perfectionnisme. L'objectif est, aussi, d'analyser les écrits de deux populations d'étudiants ayant des habiletés et des habitudes rédactionnelles contrastées en raison de la nature des tâches intellectuelles qu'ils développent régulièrement pendant leur formation universitaire. Mais surtout, il nous semble important, comparativement aux résultats des études antérieures, d'analyser plus précisément l'usage du lexique émotionnel fait par ces deux groupes d'étudiants. En effet, outre leurs habiletés rédactionnelles, ces deux groupes de rédacteurs peuvent avoir des pratiques lexicales différentes selon leurs caractéristiques de personnalité et selon l'orientation thématique de l'écrit demandé (description ou non de ses émotions et de ses sentiments).

1.1. Appréhension à écrire

D'après Daly et Miller (1975b), l'appréhension à écrire qualifie une réticence à communiquer par écrit. Pour Faigley, Daly et Witte (1981), elle consisterait à ressentir un haut niveau d'anxiété face à la nécessité de composer. Elle peut s'étendre à toute forme d'anxiété, peur, angoisse, à l'idée de rédiger (Daly, 1985). Pour l'étudier, Daly et Miller (1975a) ont conçu un instrument, permettant de la mesurer, intitulé «Writing Appréhension Test» (WAT). Daly et Hailey (1984) postulent qu'il existerait deux formes d'appréhension à écrire, l'une dispositionnelle (Trait) et l'autre situationnelle (Etat). Ainsi, la forme situationnelle (ETAT) de l'appréhension à écrire correspondrait aux émotions et sentiments transitoires très spécifiques à une situation (ou un événement donné). Elle peut changer en fonction de la tâche et de l'environnement d'écriture. La forme dispositionnelle (TRAIT)

correspondrait aux sentiments et opinions maintenus de façon relativement stable et cohérente à travers le temps et à travers les différentes situations d'écriture. C'est cette forme d'anxiété qui est mesurée par le WAT (pour une revue de question sur les mesures de l'appréhension à écrire, voir Bannour, 2007). D'après Thompson (1983), l'appréhension à écrire et l'anxiété générale seraient corrélées. Dans cette recherche, afin de disposer de mesures diversifiées de l'appréhension et même de l'anxiété qu'elle soit ETAT ou TRAIT, les deux formats du STAI ont été exploités (STAI; Bruchon-Schweitzer & Paulhan, 1993).

Plusieurs observations ont été faites à propos du comportement et du vécu subjectif des rédacteurs appréhensifs (tels que le WAT permet de les repérer) dans différents contextes. Par exemple, pour Daly et Shamo (1976, 1978), les rédacteurs les plus appréhensifs évitent les études et les carrières nécessitant une activité d'écriture (ex: études en lettres, carrières de journalisme...). Pour tenir compte de ce constat, une population d'étudiants en littérature (en principe plus habiles et motivés pour l'écriture) et une population d'étudiants en sciences (en principe peu sollicités dans leurs études à développer leurs pensées par écrit) ont été retenues pour la recherche présentée ici. À ce jour, nous ne disposons pas d'étude sur la nature du langage (syntaxique, lexicale, thématique) utilisé par des rédacteurs appréhensifs français. En revanche, aux Etats-Unis, Pennebaker et son équipe (Pennebaker, Mehl & Niederhoffer, 2003) travaillent sur cette question depuis une trentaine d'années et ont analysé les relations entre le contenu des textes et les caractéristiques émotionnelles des personnes. Ils ont procédé à de nombreuses analyses de l'écriture expressive, qui invite les rédacteurs à s'exprimer au sujet de leurs émotions et sentiments à propos d'un événement, comparativement à l'analyse d'une écriture non expressive (parler de sujets triviaux sans références aux émotions). Ils ont constaté que les procédés langagiers ne sont pas les mêmes en fonction des thématiques de productions, mais surtout en fonction de variables émotionnelles comme l'anxiété, la dépression ou l'alexithymie (NB: Ce trouble désigne un mode particulier de gestion de la vie émotionnelle se traduisant par des difficultés dans l'identification et l'expression verbale des émotions).

En ce qui concerne plus précisément la peur d'écrire, Faigley, Daly et Witte (1981) ont montré que les rédacteurs appréhensifs ont plus de difficulté à rédiger lorsque le thème concerne la description et la narration d'une expérience personnelle (exemple: 'Ecrivez à propos d'une expérience personnelle vécue qui vous a fait changer votre comportement'). En revanche, ces étudiants éprouvent nettement moins de difficulté lorsqu'ils rédigent sur un thème argumentatif n'impliquant pas l'évocation d'émotion et de sentiment (exemple: 'Argumentez pour ou contre l'idée de l'utilisation de tests psychologiques afin d'évaluer des candidats pour le diplôme de master'). Daly et Miller (1975b, 1975c), ont, de plus, montré que le volume et la qualité des écrits de rédacteurs appréhensifs étaient plus faibles que celles de rédacteurs moins appréhensifs. Toutefois, Madigan, Linton et Johnson (1996), font des observations contradictoires. Selon eux, les essais produits par les deux types de rédacteurs sont équivalents en termes de temps d'accès au lexique, de qualité de l'écrit et de complexité syntaxique. Ce sont plutôt les modèles de croyances, le sentiment d'auto-efficacité et les pensées intérieures bloquantes qui différencieraient ces deux groupes de rédacteurs. En raison de ces résultats contradictoires et de l'absence de données sur ces questions en langue française, dans la présente recherche, les rédacteurs seront incités à rédiger à partir d'une thématique expressive ou non. Le lexique émotionnel qu'ils auront utilisé sera quantifié et comparé.

1.2. Procrastination et perfectionnisme

Cette recherche s'intéresse aussi à d'autres dispositions cognitivo-émotionnelles perturbant l'activité rédactionnelle, notamment la procrastination et le perfectionnisme.

D'après Boice (1985, 1996) la procrastination traduit chez un rédacteur la tendance à remettre à plus tard la réalisation d'une tâche d'écriture malgré son intention initiale de l'effectuer dans les délais prévus. Pour, Fritzsche, Young et Hickson (2003), cette tendance à procrastiner à propos de tâches d'écriture serait associée à de l'anxiété. Plus précisément, selon Chu et Choi (2005), il existerait deux types de procrastinateurs : les procrastinateurs passifs et les procrastinateurs actifs. Les premiers seraient bloqués par leurs indécisions à agir et échoueraient par conséquent pour accomplir les tâches dans les temps. En revanche, les seconds, type plus «positif», préfèrent travailler sous la pression temporelle, et choisiraient délibérément de procrastiner pour être contraint par la nécessité impérieuse d'achever la tâche demandée. Cette distinction cruciale permet d'envisager la procrastination comme une stratégie alors adaptée à l'objectif visé. Cette stratégie relèverait d'une gestion du temps particulière associée à un sentiment d'auto efficacité: faire vite et bien au dernier moment. Lay (1986) a mis au point une échelle de mesure de la procrastination intitulée «*General Procrastination Scale*» (GPS) qui sera exploitée dans cette étude. Il sera ainsi possible de chercher à repérer les possibles relations entre appréhension à écrire et procrastination chez les étudiants ayant rédigé dans le cadre de cette recherche.

Par ailleurs, Boice (1985) définit le perfectionnisme comme l'établissement par le rédacteur d'exigences excessivement élevées, accompagnées d'autocritiques récurrentes et d'autoévaluations négatives de son écrit. Terry-Short, Owens, Slade et Dewey (1995) conçoivent une forme plus positive du perfectionnisme, et proposent de considérer cette disposition selon la fonction visée par l'individu. Cette fonction peut s'accompagner d'un perfectionnisme positif. Les cognitions et comportements de l'individu visent alors à l'atteinte d'un but associé à des conséquences positives. Elle peut aussi s'accompagner d'un perfectionnisme négatif. Les cognitions et comportements de l'individu visent l'évitement de conséquences négatives. Le perfectionnisme peut aussi être décrit selon sa source. Il peut, ainsi, être orienté vers Soi: la personne s'impose constamment des exigences élevées d'accomplissement. Il peut aussi être socialement prescrit: la personne croit que les autres ont ou entretiennent envers elle des attentes et des exigences élevées. Terry-Short et ses collègues (1995) ont développé un outil de mesure intitulé «*Positive and Negative Perfectionism Scale*» (EPPN). Par ailleurs, Onwuegbuzie (2000) a montré que des étudiants américains en doctorat avaient davantage de mauvais résultats à leurs examens et des notes moins performantes à leurs essais lorsqu'ils étaient procrastinateurs et perfectionnistes.

Au total, dans le cadre de la présente recherche, la question est de savoir si des étudiants de cursus différents (Lettres ou Sciences) présentent des niveaux de procrastination et de perfectionnisme différents et si ces niveaux influent sur la façon dont ils exploitent le lexique émotionnel dans les écrits demandés. Ces différentes observations devraient finalement permettre de mieux comprendre quelle est la nature de l'appréhension à écrire ressentie chez certains d'entre eux.

2. Méthode

2.1. Participants

60 étudiants issus d'un cursus Littéraire (lettres modernes et classiques) et 64 étudiants d'un cursus Scientifique (mathématiques, informatique et physique-chimie), ont passé l'expérience de façon collective, cursus par cursus. Ces 124 étudiants de l'Université de Provence (Aix-Marseille I) inscrits en Licence étaient volontaires et ont été répartis de façon aléatoire en 3 groupes de taille comparable. Chacun d'eux a rédigé sur un thème impliquant soit l'expression d'émotions positives, soit l'expression d'émotions négatives, soit l'absence d'expression d'émotions. Cette thématique a été attribuée de façon aléatoire à chacun des 3 groupes d'étudiants littéraires et scientifiques.

2.2. Matériel

Évaluation des caractéristiques personnelles des rédacteurs

Afin d'évaluer le caractère anxieux des participants, l'Inventaire d'Anxiété Trait-État (STAI; Bruchon-Schweitzer & Paulhan, 1993) a été utilisé. Il permet de quantifier de façon indépendante l'anxiété présente au moment de la passation (anxiété-état) et le tempérament anxieux habituel de la personne (anxiété-trait). Dans le cadre de la présente étude, le STAI état est utilisé comme un indice éventuel d'appréhension à écrire «état». Ces deux tests comportent 20 items. La personne doit répondre pour chaque item avec une échelle de Lickert en 4 points. Pour l'anxiété-état, elle évalue des items du type «Je me sens nerveux, irritable» en répondant 1=NON; 2= Plutôt NON; 3= Plutôt OUI; 4= OUI et pour l'anxiété-trait, elle évalue des items du type «J'ai un sentiment d'échec» en répondant 1= Presque jamais; 2= Parfois; 3= Souvent; 4= Presque toujours.

La version étudiante de l'Échelle Générale de Procrastination (GPS; Lay, 1986) a été utilisée. La GPS mesure de la procrastination pour 20 items au moyen d'une échelle de Lickert en 5 points. La personne indique son degré d'accord sur l'échelle (1= «Pas du tout caractéristique» à «5= Très caractéristique»). Cette échelle mesure la tendance à procrastiner lors de l'exécution de tâche (exemple: «Je ne me contrains pas à faire des tâches avant que cela me soit imposé»).

L'Échelle de Perfectionnisme Positif et Négatif (EPPN; Seidah, Bouffard & Vezeau, 2002) a été utilisée. Cette échelle est composée de 40 items regroupés en 4 sous échelles de 10 items chacune. Deux de ces échelles mesurent le perfectionnisme positif selon qu'il est personnel (*exemple*: «Je ressens de la fierté à faire les choses de façon méticuleuse») ou socialement prescrit (*exemple*: «J'aime les honneurs que me procurent mes performances remarquables»), alors que les deux autres mesurent le perfectionnisme négatif, qu'il soit personnel (*exemple*: «Quand j'entreprends quelque chose, je suis anxieux(se) à l'idée que je pourrais échouer») ou socialement prescrit (*exemple*: «J'essaie à tout prix d'éviter la désapprobation des autres»). Pour chaque item, les participants indiquent leur degré d'accord sur une échelle en 7 points (1 = tout à fait d'accord ; 6 = tout à fait en désaccord ; 7= Ne s'applique pas [réponse neutre]).

Enfin, dans le but d'identifier le niveau d'appréhension à écrire, le «Writing Apprehension Test» (WAT) de Daly et Miller (1975a) a été utilisé. Cette échelle, qui permet de mesurer l'anxiété générale ressentie à propos du fait d'écrire, est constituée de 26 items, dont 13 sont positifs (*exemple*: «J'aime que mes amis lisent ce que j'ai écrit») et 13 sont négatifs (*exemple*: «J'ai peur d'écrire»). La personne indique son degré d'accord sur une échelle en 5 points (1 = Fortement d'accord ; 5 = Fortement pas d'accord).

Analyse automatique du contenu

Afin d'étudier le lexique émotionnel intégré par les étudiants dans leur écrit, le logiciel Tropes-EMOTAIX (Tropes version 6.0; Molette, 2002; pour le scénario EMOTAIX cf. Bannour, Piolat, & Gombert, sous presse). Ce logiciel identifie automatiquement dans un texte les éléments étudiés. Plus de mille termes (qu'ils s'agissent de substantifs, de verbes, d'adjectifs, d'adverbes) sont repérés. Le lexique émotionnel étudié couvre les émotions d'arrière-plan (énergie, malaise, excitation, etc.), les émotions primaires (peur, colère, dégoût, tristesse, etc.), les émotions sociales (sympathie, embarras, honte, etc.) et les sentiments (amour, haine, etc.). Le logiciel repère aussi la valence positive (agréable) ou négative (désagréable) de ces termes.

Les textes produits par les étudiants dans cette recherche comportent en moyenne 250 mots. Avec le logiciel EMOTAIX nous avons identifié et compté l'usage des termes émotionnels selon leur valence pour chacun de ces textes. À titre d'exemple, dans l'extrait

ci-après d'un des textes, sont surlignés les termes de lexique émotionnel avec leur valence mise en expositif : Positive¹ / Négative² / Bivalent³

Consigne: Exprimez par écrit ce que vous avez ressenti lorsque vous avez obtenu votre BAC. Détaillez vos émotions, vos sentiments.

Lorsque j'ai obtenu mon bac, j'ai éprouvé³ deux sentiments³. De la joie¹ et du soulagement¹. De la joie¹ parce que je vois mon travail de l'année récompensé. Et du soulagement¹, parce que j'avais stressé² pendant l'attente des résultats.

La joie¹ que j'ai ressentie³ était grande, je me suis sentie³ très bien, de très bonne humeur¹. J'avais l'impression³ que plus rien ne pouvait m'atteindre pour briser ce moment de joie¹. J'allais pouvoir passer des vacances tranquilles. J'étais aussi heureuse¹ que mes parents car eux aussi avaient stressé² pour moi.

Comptage: Unités de lexique repérées: 8 Positives¹/ 2 Négatives²/ 5 Bivalents³. Les termes bivalents sont, dans la suite de l'analyse, qualifiés de positifs ou négatifs selon le contexte dans lequel ils sont insérés. Par exemple : «éprouvé» et «sentiments» ont été comptabilisés ici avec une valence positive.

2.3. Procédure

Pendant près d'une heure, les participants ont accompli 7 tâches dans l'ordre ci-après.

(1) Tout d'abord, ils ont répondu au questionnaire du STAI trait; (2) Puis au questionnaire du STAI état; (3) Ensuite, ils ont découvert leur thème de production écrit en haut de la page d'écriture (deux pages vierges étaient disponibles). Ils devaient écrire pendant 25 minutes. Les trois consignes concernaient les examens. L'une n'incitait pas à l'expression d'émotion : «Exprimez par écrit comment vous vous êtes préparé à passer le BAC. Détaillez votre façon de travailler et d'apprendre pendant les 15 jours qui ont précédé le BAC». Les deux autres, en tant que consignes d'écriture expressive, incitaient soit à l'expression d'émotions et sentiments positifs : «Exprimez par écrit ce que vous avez ressenti lorsque vous avez obtenu votre BAC. Détaillez vos émotions, vos sentiments», soit à l'expression d'émotions et sentiments négatifs : «Exprimez par écrit ce que vous avez ressenti lorsque vous avez échoué à un examen (BAC ou autre examen important pour vous). Détaillez vos émotions, vos sentiments »; (4) Immédiatement après cette production verbale écrite, les participants ont repassé le test du STAI état, afin de rendre compte de leur état émotionnel actuel (post-rédaction); (5) Puis, ils ont répondu au questionnaire du GPS; (6) Puis, au questionnaire de l'EPPN; (7) Enfin, ils ont répondu au questionnaire d'appréhension à écrire (WAT).

3. Résultats

Pour repérer l'effet des facteurs de personnalité sur le volume de lexique émotionnel intégré dans les textes, des analyses de régression linéaire ont été faites (N.B.: les autres analyses statistiques comme les corrélations et les Anovas ne sont pas présentées ici). Les régressions permettent de savoir si les scores des étudiants à un test donné sont (1) corrélés au volume de lexique émotionnel et si (2) ces scores «expliquent» ce volume (N.B.: Tous les résultats concernant ces analyses ne peuvent être déclinés ici, seuls les plus significatifs ont été présentés ci-après).

Il est très important de noter que, quelle que soit la consigne, pour les trois thèmes rédactionnels, les étudiants introduisent dans leur écrit des termes de lexique émotionnel de valence positive, comme de valence négative. Dans le cadre de cette publication, la présentation des résultats ne concerne que l'effet des scores de personnalité sur l'emploi du lexique émotionnel selon sa valence et selon le cursus des étudiants.

Étudiants littéraires

En consigne d'expression d'émotions positives (ressenties à l'occasion de l'obtention du BAC), les résultats sont les suivants: Les scores au Perfectionnisme Positif Socialement Prescrit et au Perfectionnisme Positif Orienté vers Soi constituent des prédicteurs significatifs ($\beta = .594, p < .004$), expliquant 40% de la variance totale de l'emploi de lexique émotionnel à valence positive [$F(2,17) = 7,45, p < .004$].

En consigne d'expression d'émotions négatives (ressenties à l'occasion de l'échec à un examen comme le BAC), les résultats sont les suivants: Le score des étudiants au Perfectionnisme Positif Orienté vers Soi a un effet faiblement significatif ($\beta = -.413$, $p < .07$), expliquant 12% de la variance totale du lexique émotionnel à valence négative [$F(1,18) = 3,70$, $p < .07$]. De plus, le score au Perfectionnisme Positif Orienté vers Soi constitue un prédicteur faiblement significatif ($\beta = .417$, $p < .06$) expliquant 12% de la variance totale du lexique émotionnel à valence positive [$F(1,18) = 3,79$, $p < .06$].

Enfin quand la consigne n'implique pas l'expression d'émotions (décrire la préparation à un examen), le score de procrastination a un effet significatif ($\beta = -.436$, $p < .05$) expliquant 14% de la variance totale de l'usage du lexique émotionnel à valence négative [$F(1,18) = 4,24$, $p < .05$].

Il faut noter que le score d'ensemble au WAT n'explique pas significativement le volume de lexique émotionnel exploité par les étudiants littéraires.

Étudiants scientifiques

En consigne d'expression d'émotions positives (ressenties à l'occasion de l'obtention du BAC), les résultats sont les suivants : le score au test anxiété «état», prédicteur significatif ($\beta = .590$, $p < .006$), explique 31% de la variance totale du lexique émotionnel à valence négative [$F(1,18) = 9,63$, $p < .006$]. Il constitue aussi, un prédicteur significatif ($\beta = -.506$, $p < .02$) expliquant 21% de la variance totale de l'emploi de lexique émotionnel à valence positive [$F(1,18) = 6,19$, $p < .02$].

En consigne d'expression d'émotions négatives (ressenties à l'occasion de l'échec à un examen comme le BAC), les résultats sont les suivants: Le score de l'anxiété état, prédicteur significatif ($\beta = .505$, $p < .01$), explique 22% de la variance totale du lexique émotionnel à valence positive [$F(1,22) = 7,57$, $p < .01$]. De plus, le score d'appréhension à écrire constitue un prédicteur significatif ($\beta = -.390$, $p < .05$) expliquant 11% de la variance totale du lexique émotionnel à valence positive employé [$F(1,22) = 5,95$, $p < .05$].

Enfin, lorsque la thématique n'implique pas l'expression d'émotions, le score d'anxiété état a un effet significatif ($\beta = -.457$, $p < .01$) expliquant 16% de la variance totale du lexique émotionnel à valence positive [$F(1,18) = 4,76$, $p < .01$].

4. Discussion

Les résultats présentés ci-avant mettent en évidence que, soumis aux mêmes thématiques rédactionnelles, les étudiants littéraires et scientifiques n'exploitent pas le lexique émotionnel de façon comparable, et ce en fonction de leurs caractéristiques de personnalités. Les facteurs dispositionnels de procrastination et de perfectionnisme sont, en effet, plus influents chez les étudiants littéraires. Ceux d'anxiété et d'appréhension à écrire le sont pour les étudiants scientifiques. Ainsi, l'ajustement à une consigne d'écriture de ces deux groupes d'étudiants relèverait d'une stratégie sans doute dépendante de leurs pratiques rédactionnelles habituelles, mais aussi de leurs caractéristiques personnelles.

Plus précisément, les étudiants littéraires ont des niveaux d'appréhension à écrire et d'anxiété «état» qui ne modulent pas significativement cette stratégie. En revanche, interviennent leur niveau de gestion du temps (procrastination), leurs exigences vis-à-vis de l'écriture ainsi que leur nécessité de sentir une auto efficacité (perfectionnisme). Au moins ils ont tendance à différer la réalisation de leurs écrits, au plus ils introduisent du lexique émotionnel dans leur écrit, s'avérant ainsi efficaces comme la consigne rédactionnelle le leur demande. Plus ils sont perfectionnistes de type positif (recherche d'une satisfaction personnelle), plus ils introduisent dans leur texte le lexique émotionnel demandé. Ces résultats sont compatibles avec la conception de Selon Boice (1985, 1996), pour qui le perfectionnisme et la procrastination (mais également les autres sources de blocage de

l'écriture évoquées plus haut) sont des composantes affectivo-cognitives d'un «schéma de pensée». D'après cet auteur, selon sa nature, le monologue interne du rédacteur – survenant avant, pendant et après une activité d'écriture – peut provoquer soit un blocage (ruminations dites «bloquantes»), soit une dynamique positive (ruminations dites «psych-up»). Ainsi, chez les rédacteurs littéraires la bonne gestion de leur temps et leur rigueur vis-à-vis de la tâche d'écriture reposeraient sur des pensées et des émotions positives. Pour ces rédacteurs, le fait d'écriture serait associé à un sentiment d'auto-efficacité, un plaisir à écrire et une satisfaction affective provoquée par l'accomplissement de la tâche.

Les étudiants scientifiques gèrent leur activité écrite en fonction de leur niveau d'appréhension à écrire et de leur anxiété «état». Plus ces étudiants sont appréhensifs à écrire, plus ils sont anxieux, et plus ils introduisent de lexique émotionnel dans leur écrit. Tout se passe comme si leurs caractéristiques de personnalité agissaient sur leur éveil affectif et cognitif en leur permettant de répondre efficacement à la consigne rédactionnelle. Autrement dit, leurs dispositions émotionnelles ne les pénalisent pas, comme certains travaux sur l'appréhension à écrire pouvaient le laisser penser (Daly, 1985). Au contraire, plus ils ont de l'anxiété, plus ils exploitent cette anxiété pour réaliser la tâche rédactionnelle demandée. Néanmoins, l'engagement dans l'écriture s'opère sur des bases de nature anxieuses. La difficulté ainsi ressentie permet certes de stimuler l'éveil cognitif du rédacteur, mais provoque également un sentiment de malaise lors de la réalisation de produits écrits. Il en découle, au fil des années, et comme l'ont montré les résultats aux échelles de mesure des composantes affectivo-émotionnelles, un ressenti d'incompétence et un dégoût de l'écriture.

En conclusion, cette étude laisse à penser que pour les étudiants qui ont participé à cette expérience, le processus rédactionnel s'accompagne d'émotions et de pensées positives pour les étudiants littéraires et d'émotions et de pensées négatives pour les scientifiques. Ce constat doit être, évidemment, précisé et confirmé par d'autres recherches. Par ailleurs, afin de mieux comprendre l'impact des dimensions cognitivo-émotionnelles sur le contenu des écrits, il faudra aussi travailler sur l'usage du lexique émotionnel dans d'autres types d'écrits (essais, rapports, poèmes, blogs). Les observations présentées ici ne constituent qu'une première étape de l'évaluation des usages lexicaux de rédacteurs adultes en fonction de leur personnalité.

Références bibliographiques

- BANNOUR, Rachid (2007). Outils de mesure des composantes émotionnelles de l'appréhension à écrire. *Sciences Croisées*.
[http://www.up.univ-mrs.fr/divrech/ed356_id/index.htm]
- BANNOUR, Rachid, PIOLAT, Annie, & GOMBERT, Anne (sous presse). Quand des Lycéens et des étudiants rendent compte de leurs émotions... Effets de leur niveau d'appréhension à écrire et du thème rédactionnel, in *Lexique et production verbale : Vers une meilleure intégration des apprentissages lexicaux*, Grossmann, Francis & Plane, Sylvie (Eds). Villeneuve d'Ascq, Presses Universitaires du Septentrion.
- BOICE, Robert (1985). Cognitive components of blocking. *Written Communication*, 2: 91-104.
- BOICE, Robert (1996). *Procrastination and blocking: A novel, practical approach*. Praeger Publishers/Greenwood Publishing Group.
- BRUCHON-SCHWEITZER, Marilou, & PAULHAN, Isabelle. (1993). *Adaptation française de l'Inventaire d'Anxiété Trait-Etat. Forme Y (STAI Y)*. Paris: Les Éditions du Centre de Psychologie Appliquée.
- CHU, Angela Hsin Chun, & CHOI, Jin Nam (2005). Rethinking Procrastination: Positive Effects of "Active" Procrastination Behavior on Attitudes and Performance. *The Journal of Social Psychology*, 145(3): 245-264

- DALY, John, & MILLER, Michael (1975a). The empirical development of an instruction to measure writing apprehension. *Research in the Teaching of English*, 9: 242-249.
- DALY, John, & MILLER, Michael (1975b). Apprehension of writing as a predictor of message intensity. *Journal of Psychology*, 89: 175-177.
- DALY, John, & MILLER, Michael (1975c). Further attitudes in writing apprehension: SAT scores, success expectations, willingness to take advanced courses and sex differences. *Research in the Teaching of English*, 9: 250-256.
- DALY, John, & HAILEY, Joy-Lynn (1984). Putting the situation into writing research: Situational parameters of writing apprehension as disposition and state, in *New directions in composition research*, Beach, Richard & Bidwell, Lillian (Eds). New York: Guilford, 259-274.
- DALY, John, (1985). Writing apprehension, in *When a writer can't write: Studies in writer's block and other composing process problems*, Rose, Mike (Ed) New York: Guilford, 43-82.
- DALY, John, & SHAMO, Wayne (1976). Writing apprehension and occupational choice. *Journal of Occupational Psychology*, 49: 55-56.
- DALY, John, & SHAMO, Wayne (1978). Academic decisions as a function of writing apprehension. *Research in the Teaching of English*, 12: 119-126.
- FAIGLEY, Lester., DALY, John, & WITTE, Stephen (1981). The role of writing apprehension in writing performance and writing competence. *Journal of Educational Research*, 7:, 16-21.
- FRITZSCHE, Barbara, YOUNG, Beth Rapp, & HICKSON, Kara (2003). Individual differences in academic procrastination tendency and writing success. *Personality and individual differences*, 35: 1549-1557.
- HAYES, John (1996). Un nouveau cadre pour intégrer cognition et affect dans la rédaction de texte, in *La rédaction de textes*, Piolat, Annie & Pélissier, Aline (Eds). Lausanne : Delachaux et Niestlé, 51-101.
- LAY, Clarry (1986) At last, my research on Procrastination. *Journal of Research in Personality* 20:474-495.
- MADIGAN, Robert, LINTON, Patricia & JOHNSON, Susan (1996). The paradox of writing apprehension, in *The science of writing, Theories, methods, individual differences, and applications*, Levy, Michael & Ransdell, Sarah (Eds). Mahwah (New Jersey): Lawrence Erlbaum Associates, 295-307.
- MOLETTE, Pierre (2002). Tropes. ACETIC (Action Études Information Communication). CD ROM.
- OLIVE, Thierry & PIOLAT, Annie (2003). Activation des processus rédactionnels et qualité des textes. *Le Langage et l'Homme*, 28(2): 191-206.
- ONWUEGBUZIE, Anthony (2000). Academic Procrastinators and Perfectionistic Tendencies Among Graduate Students (special issues). *Journal of Social Behavior and Personality*, 15: 103-110.
- PENNEBAKER, James, MEHL, Matthias, & NIEDERHOFFER, Kate. (2003) Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, 54 : 547-577.
- SEIDAH, Amélie, BOUFFARD, Thérèse, & VEZEAU, Carole (2002). La mesure du perfectionnisme: Validation canadienne-française du Positive and Negative Perfectionism Scale. *Revue canadienne des sciences du comportement*, 34: 168-171.
- TERRY-SHORT, L.A, OWENS, Glynn, SLADE, Peter, & DEWEY, Michael (1995). Positive and negative perfectionism. *Personality and Individual Differences*, 18: 663-668.
- THOMPSON, Merle O'Rourke. (1983). The returning student: Writing anxiety and general anxiety. *Teaching English in the Two-Year College*, 10(1): 35-39.

UNE DIFFICULTÉ DE LA DESCRIPTION LINGUISTIQUE : L'OBJET DIRECT EN FRANÇAIS

Florinela COMANESCU

florinela_comanescu@yahoo.fr

Université de Bucarest

36 – 46, Boulevard Mihail Kogalniceanu

70609 Bucarest

Roumanie

Université de Paris 7 Denis Diderot (LATTICE)

2, Place Jussieu

75005 Paris

Abstract : Studies written on this topic are discussed and a hierarchization of the classical tests used for the direct object in French is proposed. We argue that the definition of this constituent should be based on the pronominalization test, the other tests being used to describe it. This paper also identifies certain aspects which have been insufficiently studied.

1. Préliminaires

Les difficultés empiriques liées à la délimitation/description du domaine recouvert par la notion d'objet direct sont bien connues : la décision radicale de M. Gross à ce sujet (Gross, 1969 : 72-73) en donne pleinement la mesure. A ces difficultés d'ordre descriptif s'ajoutent celles relevant de la naissance relativement récente de la notion et les divergences terminologiques d'une langue à l'autre ou d'un modèle théorique à l'autre (Willems, Melis, 1998 : 5-6).

Pour la plupart des linguistes, la nécessité de la notion d'objet l'emporte cependant sur les difficultés qu'elle soulève, sans qu'aucune solution définitive ait pu être proposée pour le moment et malgré le nombre impressionnant d'études consacrées à ce sujet.

Nous proposons dans ce travail une discussion de la notion syntaxique d'objet direct en français du point de vue des critères utilisés pour la délimitation/description du domaine qu'elle recouvre, ainsi qu'un bilan des nombreuses études qui portent sur ce sujet.

Notre travail a comme point de départ deux idées récurrentes dans les études : la nécessité d'une description multifactorielle de ce phénomène, dont découle également son caractère scalaire et l'idée que la transitivité n'engage pas uniquement le verbe, mais l'énoncé dans son ensemble. Notre propre contribution consistera à proposer une hiérarchisation des critères de la transitivité, selon le niveau d'incidence de la phrase/énoncé auquel ils s'appliquent.

2. La définition multifactorielle de la transitivité

2.1. *Types d'approches*

La définition de la transitivité prend en considération une multitude de critères : on parle à ce sujet de « faisceaux de propriétés différentes pour définir l'Objet ; le point décisif étant en quelque sorte que l'Objet soit une composition de propriétés. » (Bortolussi, 1990 : 30), de « scintillement de propriétés associées à cette notion, qui ne peut être capté qu'à l'aide d'un éventail de traits riche et diversifié. » (Abeillé, 1997 : 8), ou tout simplement de « théorie multifactorielle de la transitivité » (François, 1998).

L'inventaire des tests utilisés dans l'approche de l'objet direct est relativement facile à établir à travers les études qui portent sur ce sujet, mais non l'organisation du système qu'ils forment, les objectifs des recherches n'étant que rarement convergents. Généralement, les tests de l'objet sont envisagés comme des instruments de découverte vérifiés et donnés d'emblée, dont il n'est plus nécessaire ou utile de discuter la pertinence. Leur rôle consiste à mesurer des différences/décalages syntaxiques et sémantiques au niveau du fonctionnement des constructions, mais on le fait, paradoxalement, sans avoir vérifié préalablement ce que chacun de ces tests mesure effectivement. Les réponses convergentes par rapport aux tests de la transitivité ne prouvent pas du tout l'équivalence des tests eux-mêmes, mais tout simplement l'existence dans la langue de verbes/constructions qui relèvent de ce que l'on appelle la « transitivité prototypique » (Desclés, 1998 : 161-180). C'est le cas, par exemple, des 650 verbes sélectionnés par D. Willems (Willems, 1981 : 200-207), qui ne figurent qu'à l'intérieur de la construction transitive directe, et pour lesquels les tests classiques de l'objet reconnus par l'auteur (pronominalisation, transformation passive, emploi absolu) ne présentent pas de valeur opératoire.

2.2. Inventaire des tests

Les inventaires de tests utilisés dans les études présentent des variations. La liste la plus complète est celle systématisée par L. Melis (Melis, 2001 : 254-246) : propriétés de codage (l'existence d'indices clitiques – *le, la, les, en*, la position postverbale comme position privilégiée, la coalescence étroite avec le verbe) et propriétés de fonctionnement (la capacité à fonctionner comme sujet dans les reformulations passives, périphrastiques ou réflexives, la saturation de la position d'objet dans la construction factitive, la capacité à servir de pivot pour l'infinitif construit avec *à* dans diverses constructions : *GN être Adj. à Vinf*, *GN (être) à Vinf*, *il y a GN à Vinf*).

Un autre inventaire assez complet est proposé par L. Pino Serrano (Pino Serrano, 2000 : 411-417). La distinction entre critères (la construction directe, la place, le caractère nucléaire) et tests (l'interrogation, la pronominalisation, la passivation, l'accord du participe passé) proposée par cet auteur rejoint partiellement celle opérée par L. Melis.

Chez d'autres linguistes, les inventaires de tests sont généralement plus réduits. Il arrive même que la transitivité des verbes/constructions soit définie par rapport à un test fonctionnel unique, tel la passivation (à côté cependant de deux types de phénomènes non-fonctionnels : le glissement sémantique et les faits d'ordre) (Rousseau, 1998 : 92).

Bien que les critères de la transitivité soient traités comme relevant du même plan et ayant le même degré de pertinence, ils font cependant dans les études l'objet d'une hiérarchisation implicite. Trois tests notamment sont concernés par cette hiérarchisation : la pronominalisation, la transformation passive et l'emploi absolu.

Nous pensons d'ailleurs que reconnaître la théorie scalaire de la transitivité (Lazard, 1998 : 55-84, Desclés, 1998 : 161-180) revient en fait à accepter également la scalarité des tests eux-mêmes, qui eux seuls permettent de discerner des différences entre les types d'objets. « L'inévitable et nécessaire hiérarchie des objets » (Rousseau, 1998 : 109) ne nous paraît nullement possible en l'absence d'une hiérarchisation complémentaire des tests qui dégagent ces premiers.

3. La transitivité comme propriété de l'énoncé

L'idée que la transitivité ne relève pas du verbe, mais de l'énoncé dans son ensemble est récurrente dans les études : « [...] il n'y a pas grand intérêt à vouloir à tout prix classer ainsi les verbes français, et il serait plus important d'expliquer la variété de leurs emplois. » (Creissels, 1995 : 248), « [...] LE VERBE TOMBE DE SON PIEDESTAL ET LES COMPLEMENTS FONT LA LOI. » (Wilmet, 1997 : 479), « [...] la transitivité n'est pas une

donnée immédiate du lexème verbal. » (Rousseau, 1998 : 96), « La transitivité est une propriété non du verbe seul, mais de l'ensemble du complexe prédicatif. » (Rousseau, 1998 : 104), « Le français contemporain vit une période de transitivité vacillante ; le manque de prédisposition actancielle – sa fluctuation – en témoigne. » (Larjavaara, 1998 : 313), bien que la théorie contraire soit également formulée : « Un verbe est donc transitif ou intransitif ; la transitivité est une question de oui ou non. » (Lazard, 1998 : 55).

Dans cette situation, il serait peut-être plus utile de renoncer à éliminer certains phénomènes et de chercher à envisager tout simplement comment ils se manifestent et dans quelle mesure ils pourraient s'avérer pertinents dans la discussion de la transitivité.

3.1. Le niveau du groupe verbal. La relation verbe-objet. La pronominalisation

Devant l'extrême variété des effets de sens attribuables à la fonction syntaxique d'objet direct, les études proposent généralement pour cette notion une définition syntaxique/fonctionnelle : « substituer autant que possible la notion de fonction à celle de sens » (Rothemberg, 1974 : 19), « [...] il nous faut admettre la nécessité d'une notion purement formelle d'objet direct. » (Gaatone, 1998a : 15)

L'élément commun dans les définitions est le test de la pronominalisation personnelle dans le cas de la détermination définie, accompagnée ou non de tests supplémentaires : la dislocation (Blinkenberg, 1960 : 127), la variabilité du déterminant et l'impossibilité de commutation du substantif postverbal déterminé avec un adjectif (Rothemberg, 1974 : 17), le *de* en contexte négatif et le quantitatif *en* (Creissels, 1995 : 233, 238-239), ou la pronominalisation adverbiale en *en* accompagnée de la possibilité du déterminant *de* indéfini induit par la négation (Abeillé, 1997 : 23).

Tous les autres critères sauf celui de la pronominalisation personnelle nous semblent cependant présenter des limites : soit ils s'appliquent à d'autres constituants également – le *en* paraît lié plutôt à la position qu'à la fonction objet (Lagae, 1998, 103-113), le *de* négatif concerne aussi le constituant nominal postverbal des constructions impersonnelles, (Gaatone, 1998a : 15-16) – soit ils présentent une portée beaucoup plus réduite que celle de la pronominalisation (la construction causative et la construction « *facile à dire* » sont opérationnelles uniquement dans les phrases possédant comme sujet un constituant humain).

Dans cette situation, deux solutions nous semblent envisageables : considérer que ces tests présentent tous le même degré de pertinence et délimiter un domaine extrêmement restreint correspondant à la notion d'objet direct, ou accepter tout simplement que ces phénomènes ne tiennent pas du même niveau de l'approche et proposer de les distinguer.

La première solution présente au moins deux inconvénients. D'une part, elle obligerait à accepter que des constituants ayant le même comportement par rapport à la pronominalisation personnelle puisse ne pas se rattacher à la fonction syntaxique d'objet direct, position qui rendrait incohérente toute tentative de description. Dans un commentaire portant justement sur la pertinence des deux formes de pronominalisation (personnelle et adverbiale), D. Gaatone (Gaatone, 1998b : 74) semble préférer la pronominalisation adverbiale. L'inconvénient de la pronominalisation personnelle consiste pour lui, justement, dans le fait qu'elle ne permet pas d'éliminer des constituants tels les compléments de mesure construits directement. Si l'intérêt de la démarche vise, au contraire, à pouvoir rendre compte des relations entre les différents types de constituants et non pas à opérer des distinctions et des éliminations tranchantes, nous pensons que la pronominalisation personnelle peut prouver pleinement sa pertinence opérationnelle.

D'autre part, en l'absence de tests formels discriminatoires (puisque tous les tests apporteraient des réponses convergentes), la description du domaine par des procédés

fonctionnels deviendrait impossible. De plus, ces tests ne pourraient pas opérer pour les phénomènes non-prototypiques, écartés dès le début, et dont on aurait nié tout rapport avec la transitivité et avec les tests de celle-ci. Par conséquent, la solution la plus convenable pour la cohérence de la description et de plus, la plus adaptée à la réalité des faits, nous semble la deuxième : les tests formels n'appartiennent pas au même plan, ils n'interviennent pas au même moment dans l'approche de l'objet et il s'impose d'opérer une hiérarchisation entre eux.

La promotion de la pronominalisation au rang de test fondamental dans la définition de l'objet direct nous apparaît parfaitement juste et elle correspond d'ailleurs à la plupart des propositions faites dans les études. Seulement aucun argument n'est fourni dans ce but, ce qui fait que cette sélection puisse risquer de paraître aléatoire.

L'argument que nous apportons à ce sujet est lié notamment à l'existence de plusieurs niveaux d'incidence des tests sur la phrase. Ainsi, la pronominalisation concerne uniquement la relation syntaxique entre le verbe et l'objet, aucun autre ordre de facteurs ne pouvant la changer. Autrement dit, une fois qu'un constituant accepte ce test, aucun facteur ne peut intervenir pour modifier sa réponse, comme c'est le cas pour les autres tests. La pronominalisation ne concerne que la relation syntaxique entre le verbe et son objet et elle n'est soumise à aucun facteur extérieur à cette relation.

3.2. Le niveau phrastique. La relation sujet-objet. La transformation passive

La phrase/transformation passive a fait l'objet d'un nombre d'études impressionnant et elle figure dans la plupart des cas parmi les critères définitoires de l'objet direct. Cependant, il est plutôt rare que la définition de l'objet direct soit fondée sur ce critère formel exclusivement (Rousseau, 1998 : 92).

A la différence de la pronominalisation, la problématique de la passivation concerne la phrase dans son ensemble. Nous ne faisons pas allusion au fait généralement invoqué (mais vrai seulement partiellement) que la passivation implique une réorganisation de tous les constituants de la phrase (ou des constituants relevant de la phrase minimale, au moins), bien que cet aspect soit important lorsqu'il s'agit de discuter le niveau concerné par la transformation passive.

Nous nous rapportons à un autre phénomène, signalé et discuté dans plusieurs travaux, qui concerne le fait que la passivation ne se décide pas au niveau du mot ou du syntagme, mais au niveau de la phrase entière. Autrement dit, la distinction entre verbes passivables et verbes non-passivables ne résout nullement le problème de la transitivité, elle délimite tout simplement la classe des verbes pour lesquels la discussion de la transformation passive est pertinente. Une fois cette classe délimitée, il faut procéder ensuite à cerner les différents facteurs de la passivation, ce qui peut impliquer une discussion pour chaque cas pris séparément. Nous fondons notre affirmation sur deux ordres de phénomènes également signalés dans les études. Il s'agit du rôle que des constituants autres que ceux directement concernés par la transformation passive peuvent avoir à déterminer la passivabilité d'une phrase, d'une part, et de l'importance de la relation sujet-objet, telle qu'elle est constituée via le verbe de la phrase, d'autre part.

Nous illustrons le premier cas par des exemples discutés par D. Gaatone, en vue de démontrer le rôle des constituants autres que ceux directement concernés par cette transformation à faire accepter le passif :

?? Une montre a été trouvée par moi.

La montre a été trouvée par moi et non par ma sœur. (Gaatone, 1998b : 262)

Ce qui rend la phrase passive possible, dans l'opinion de l'auteur, c'est la création d'un cadre contrastif avec le pronom personnel, par la présence d'un autre constituant (la présence d'un sujet pronom personnel de la première ou deuxième personne bloquant généralement la construction passive – Lamiroy, 1993 : 58, Gaatone, 1998b : 262).

Le second exemple que nous choisissons pour illustrer la dimension phrastique de la transformation passive suppose l'existence d'un « différentiel de participation » (François, 1998 : 181-201) entre les constituants sujet et objet et prend en considération notamment le contenu sémantique véhiculé par la phrase. Nous proposons le cas du verbe *rencontrer*, qui a fait l'objet de plusieurs discussions dans les études.

Ainsi, si les deux constituants du verbe *rencontrer* sont réalisés par des noms humains, la passivation est bloquée, en raison du différentiel de participation réduit : **Paul a été rencontré (fortuitement ou non) par un partenaire*. En échange, la passivation devient possible si le nom objet désigne une autre entité qu'un humain, ce qui augmente le différentiel de participation : *Beaucoup d'obstacles ont été rencontrés dans cette entreprise* (François, 1998 : 193).

La discussion proposée par P. Blumenthal (Blumenthal, 1998 : 59-60) appuie l'idée que la passivation est fonction de la relation entre le sujet et l'objet et non pas de la nature de chacun de ces constituants pris séparément. Ainsi, ses exemples contenant *rencontrer* semblent contredire dans un premier temps son hypothèse de travail : plus l'objet partage les propriétés d'un sujet syntaxique prototypique, plus il a de chances d'assumer le rôle de sujet logique, donc de faire l'objet d'une transformation passive. Dans une telle perspective, la passivation des phrases possédant deux constituants humains devrait aller de soi, mais le comportement de la phrase avec *rencontrer* dément cette hypothèse : **Mon ami a été rencontré hier*. (Blumenthal, 1998 : 59). L'auteur ne discute pas le problème de ce verbe en termes de contribution au développement du procès/participation, mais du point de vue de l'existence d'un sujet-source antérieur au procès, qui fonctionne en tant que cause de celui-ci. L'explication qu'il fournit rejoint cependant celle de J. François (François, 1998 : 181-201) : il ne suffit pas que les deux constituants soient de la même nature, il est surtout nécessaire qu'ils ne soient pas traités de la même façon. Pour l'exemple en question, cette deuxième condition n'est pas remplie, et c'est cette particularité qui bloque la passivation.

La passivation n'est donc pas une opération mécanique de permutation, elle repose sur l'existence de relations syntaxiques et sémantiques complexes entre les constituants, un changement minime pouvant fonctionner comme un facteur d'autorisation, ou au contraire, de blocage de la transformation. De même, elle dépasse le niveau du syntagme verbal et elle engage la phrase toute entière.

3.3. *Le niveau discursif. L'emploi absolu*

Bien qu'aucune étude ne fonde la définition de l'objet direct sur l'emploi absolu, ce test est régulièrement mentionné dans les études et fait également l'objet de nombreuses discussions. L'emploi absolu découle directement du caractère obligatoire du constituant objet : l'absence syntaxique d'un constituant prévu dans la valence du verbe n'entraîne pas l'élimination au niveau sémantique du participant au procès désigné, mais signale tout simplement une asymétrie entre les niveaux syntaxique et sémantique de la phrase.

Les problèmes soulevés par l'emploi absolu sont nombreux, mais nous nous arrêterons uniquement sur deux points qui nous semblent éclairants pour notre discussion : le corpus étudié et les facteurs favorisant l'emploi absolu.

Pour ce qui est du corpus discuté, ce n'est que dans peu de cas que le niveau phrastique s'avère suffisant à rendre compte de l'emploi absolu, comme dans le cas des emplois habituels (Noailly, 1998a : 130-143). Dans la plupart des cas l'observation porte au moins

sur des séquences textuelles minimales, sinon sur de vrais paragraphes. L'approche de l'emploi absolu dépasse d'ailleurs le niveau purement linguistique, pour se placer au niveau de l'évaluation des connaissances extralinguistiques des usagers de la langue. S'il se manifeste au niveau phrastique, plus exactement au niveau du syntagme verbal, son explication est à chercher au-delà de ce niveau, dans l'environnement textuel et même au niveau des paramètres extralinguistiques : « Il faut prendre en compte tout le contexte ; le verbe lui-même et le genre de procès qu'il décrit – habituel ou non – ne suffisent guère. » (Larjavaara, 2000 : 133), « Il n'y a pas de restriction lexicale : tout verbe transitif peut se trouver sans objet dans un emploi générique, si la situation extralinguistique s'y prête. » (Larjavaara, 2000 : 136).

Pour ce qui est des paramètres de l'emploi absolu, à la limite, si les conditions pour la bonne identification/récupération du référent non exprimé linguistiquement sont bien remplies, tous les verbes transitifs du français sont susceptibles de fonctionner sans objet direct : l'emploi absolu est premièrement un phénomène discursif, qui relève de la « mémoire discursive » (Noailly, 1998b : 45).

Dans une telle perspective, l'emploi absolu nous semble pouvoir être exploité dans un tout autre domaine que celui de la transitivité : si tous les verbes du français sont susceptibles de présenter ce fonctionnement (dans des contextes précis, évidemment), alors il ne peut pas servir à rendre compte de la transitivité des verbes, mais tout simplement de la pertinence du contexte/co-texte à suppléer l'absence de l'objet et à guider dans l'identification du référent correspondant au constituant qui manque. L'emploi absolu n'est plus un test de la transitivité, mais tout simplement l'une des diverses manifestations discursives de celle-ci.

3.4. *Les autres tests de la transitivité*

Pour ce qui est des autres tests formels de la transitivité (Gaatone, 1998a : 13-20), ils ne font pas l'objet de discussions particulières, même si D. Gaatone les considère comme ayant le même degré de pertinence que les tests déjà discutés. Nous souhaitons faire quelques remarques sur deux de ces tests uniquement (la construction causative et la construction « *facile à dire* »), les deux retenus par L. Melis (Melis, 2001 : 243-257) dans son inventaire.

Tout en souhaitant éviter les généralisations dangereuses, nous croyons que ces tests concernent uniquement la dimension sémantique de la transitivité : ils ne portent, à notre avis, que sur les constructions possédant comme sujet un constituant qui désigne un humain. D'ailleurs, le test avec le verbe *faire* est l'un des tests classiques de l'agentivité, les deux tests discutés nous apparaissant comme de simples variantes fonctionnelles de ce test. D. Gaatone signale le rapport étroit entre les constructions du type *Nom à Infinitif et Adjectif à Infinitif* et le problème sémantique de l'agentivité : « [...] les constructions en question sont limitées non seulement aux verbes transitifs directs, mais aussi aux verbes agentifs. » (Gaatone, 1998b : 104).

Arrivée à ce point de notre discussion, nous croyons pouvoir formuler encore une remarque concernant les études sur l'objet direct/transitivité. Bien que les définitions sémantiques de la transitivité proposées dans les grammaires appelées traditionnelles soient généralement rejetées, de par leur caractère réducteur et excessivement généralisateur, il semble cependant que la plupart des approches modernes de ce phénomène continuent à travailler sur le même genre de corpus, ce qui peut amener, inévitablement, à des conclusions du même type. Par conséquent, une part importante du domaine de la transitivité reste malheureusement inabordable pour le moment, les études portant généralement sur des constructions transitives prototypiques, caractérisées par un certain sémantisme (transformation/création) et par des paramètres syntaxiques particuliers (noms humains pour le sujet, noms concrets pour l'objet). En s'occupant des constructions non-prototypiques ayant comme objet des noms de

procès, A. Celle (Celle, 1997 : 139-172) explore justement l'une des zones d'ombre de la transitivité.

4. Conclusions

Les tests classiques de la transitivité ne relèvent pas du même plan de description et ne mesurent pas le même type de paramètres. La hiérarchisation des tests, correspondant à la hiérarchisation des plans, révèle la nécessité de plusieurs distinctions à prendre en considération : définition/description, fonctionnement (paramètre)/ingrédient de la transitivité.

Nous croyons que la définition de l'objet direct doit reposer uniquement sur le test de la pronominalisation : c'est le seul fonctionnement stable de l'objet, constant à tous les niveaux et n'ayant pas à être modifié par de possibles changements intervenus dans la phrase. Une fois le domaine de la transitivité ainsi délimité, les autres tests servent à en donner la description, avec une discussion complémentaire obligatoire des paramètres de tout type qui la déterminent.

Nous ne souhaitons cependant pas affirmer que la transitivité est un phénomène uniquement discursif, soumis aux variations et aux fluctuations de ce domaine, au contraire. Si toute variation discursive (et même phrastique peut-être) peut finalement être réglée par des réductions successives, c'est parce qu'elle est inscrite et sinon validée au moins autorisée au niveau de la langue. Qu'il s'agisse du verbe ou des types de noms objet et sujet, le premier niveau à prendre en considération est celui du mot-même, que nous avons délibérément, bien qu'injustement, à coup sûr, laissé de côté dans cette discussion, uniquement pour des raisons liées à l'économie de notre travail et au but précis et restreint que nous nous y sommes fixé.

Références

- ABEILLE, Anne (1997). Fonction ou Position Objet ? (II et fin), *Le gré des langues*, 12 : 8-33.
- BLINKENBERG, Andréas (1960). *Le problème de la transitivité en français moderne : essai syntactico-sémantique*. Copenhague : Det Kongelige Danske Videnskabernes Selskab.
- BLUMENTHAL, Peter (1998). Le complément peut-il être sujet logique, in *Prédication, assertion, information : Actes du colloque d'Uppsala en linguistique française, 6-9 juin 1996, Studia Romanica Upsaliensia* 56, Forsgren, Mats, Jonasson, Kerstin & Kronning, Hans (Eds). Uppsala : Acta Universitatis Upsaliensis, 57-66.
- BORTOLUSSI, Bernard (1990). Objet d'une théorie, *Le gré des langues*, 1 : 26-35.
- CELLE, Agnès (1997). Quand l'objet est un nom de procès, *Cahiers Charles V*, 23 : 139-172.
- CREISELS, Denis (1995). *Eléments de syntaxe générale*. Paris : Presses Universitaires de France.
- DESCLES, Jean-Pierre (1998). Transitivité sémantique, transitivité syntaxique, in *La transitivité*, Rousseau, André, (Ed). Villeneuve d'Ascq : Presses Universitaires du Septentrion, 161-180.
- FRANCOIS, Jacques (1998). Théorie multifactorielle de la transitivité, « différentiel de participation » et classes aspectuelles et actanciennes de prédication, in *La transitivité*, Rousseau, André, (Ed). Villeneuve d'Ascq : Presses Universitaires du Septentrion, 181-201.
- GAATONE, David (1998a). L'objet direct comme notion formelle dans la formulation des règles syntaxiques, in *Les objets : relations grammaticales et rôles sémantiques*, Willems, Dominique & Melis, Ludo (Eds). Bruxelles : De Boeck & Larquier, Département Duculot, 13-20.
- GAATONE, David (1998b). *Le passif en français*. Paris : Duculot.
- GROSS, Maurice (1969). Remarques sur la notion d'objet direct en français, *Langue française*, 1 : 63-73.

- LAGAE, Véronique (1998). En quantitatif : pronom lié à la fonction objet ou à une position ?, in *Les objets : relations grammaticales et rôles sémantiques*, Willems, Dominique & Melis, Ludo (Eds). Bruxelles : De Boeck & Larcier, Département Duculot, 103-113.
- LAMIROY, Béatrice (1993). Pourquoi il y a deux passifs, *Langages*, 109 : 53-72.
- LARJAVAARA, Meri (1998). Sur les variations de la transitivité en français contemporain, in *Prédication, assertion, information : Actes du colloque d'Uppsala en linguistique française, 6-9 juin 1996*, *Studia Romanica Upsaliensia* 56, Forsgren, Mats, Jonasson, Kerstin & Kronning, Hans (Eds). Uppsala : Acta Universitatis Upsaliensis, 307-315.
- LARJAVAARA, Meri (2000). *Présence ou absence de l'objet : limites du possible en français contemporain*. Helsinki : Academia Scientiarum Fennica.
- LAZARD, Gilbert (1998). De la transitivité restreinte à la transitivité généralisée, in *La transitivité*, Rousseau, André, (Ed). Villeneuve d'Ascq : Presses Universitaires du Septentrion, 55-84.
- MELIS, Ludo (2001). Les compléments nominaux des verbes de mouvement intransitifs et la constellation de l'objet, in *Par monts et par vaux*, Buridant, Claude, Kleiber, Georges & Pellat, Jean-Claude (Eds). Paris : Editions Peeters, 243-257.
- NOAILLY, Michèle (1998a). Emploi absolu, anaphore zéro et transitivité, in *La transitivité*, Rousseau, André, (Ed). Villeneuve d'Ascq : Presses Universitaires du Septentrion, 127-144.
- NOAILLY, Michèle (1998b). Les traces de l'actant objet dans l'emploi absolu, in *Les objets : relations grammaticales et rôles sémantiques*, Willems, Dominique & Melis, Ludo (Eds). Bruxelles : De Boeck & Larcier, Département Duculot, 39-47.
- PINO SERRANO, Laura (2000). Le complément d'objet direct en français : problèmes de définition et de représentation, *Actes du XXII-ème Congrès International de Linguistique et Philologie Romanes*, Bruxelles, 1998. Bruxelles : Niemeyer, 411-417.
- ROTHEMBERG, Mira (1974). *Les verbes à la fois transitifs et intransitifs en français contemporain*. Paris – The Hague : Mouton.
- ROUSSEAU, André (1998). La double transitivité existe-t-elle ? Réflexions sur la nature de la transitivité, in *La transitivité*, Rousseau, André, (Ed). Villeneuve d'Ascq : Presses Universitaires du Septentrion, 85-112.
- WILLEMS, Dominique (1981). *Syntaxe, lexique et sémantique. Les constructions verbales*. Gent : Rijksuniversiteit te Gent.
- WILLEMS, Dominique, MELIS, Ludo, (1998). Introduction, in *Les objets : relations grammaticales et rôles sémantiques*, Willems, Dominique & Melis, Ludo (Eds). Bruxelles : De Boeck & Larcier, Département Duculot, 5-10.
- WILMET, Marc (1997). *Grammaire critique du Français*. Paris : Duculot.

STRATÉGIES DE RESTITUTION DES CONSTRUCTIONS VERBALES SÉRIELLES DU CRÉOLE HAÏTIEN EN FRANÇAIS L2

Darline COTHIÈRE-ROBERT

dacoty25@yahoo.fr

Université Paris 3 – Sorbonne nouvelle /EA 2288 DILTEC

45, rue Saint-Jacques

75005 Paris

Abstract : This article compares Haitian Creole to French in terms of the use of serial verbal constructions, a grammatical structure not recognized in French. Our main focus will be on strategies used to keep the use of such verbal constructions when learning French as a second language.

1. Introduction

Cet article propose une comparaison du créole haïtien au français à partir des constructions verbales sérielles, une structure grammaticale qui n'est pas attestée en français. Nous analysons ici les différentes stratégies de restitution de ces structures en français L2 à travers une activité traductionnelle par un groupe de trente créolophones haïtiens. L'objet est de savoir si ces constructions imposent des contraintes syntaxiques et/ou sémantiques dans le passage de L1 à L2 et si cette activité donne lieu à des phénomènes de transfert récurrents liés à la spécificité structurelle de la L1.

2. Les constructions verbales sérielles : problèmes, définition et exemples en créole haïtien

Les constructions verbales sérielles (désormais CVS) sont des propositions à prédicat complexe, attestées notamment dans des langues de l'Afrique de l'Ouest (Bole-Richard, 1978 ; Givón, 1991) dans la langue chinoise (Hagège, 1975) et dans plusieurs langues créoles (Valdman, 1978 ; Manessy, 1995). Si déterminer la filiation des CVS dans les langues créoles demeure problématique¹, les catégoriser linguistiquement l'est tout autant. Certains auteurs y voient une catégorie grammaticale (Bole-Richard, 1978), certains autres, une catégorie lexicale (Valdman, 1978). Il existe par ailleurs des positions partagées en ce qui a trait à la portée sémantique voire cognitive de ce type de structure. Il s'agirait, dans un cas, d'un mécanisme de décomposition analytique du procès où les composants du signifié (Sé) sont exprimés par des signifiants (Sa) verbaux distincts (Manessy, 1995) ; dans un autre cas, d'un mécanisme d'intégration du procès en vue d'une opération déterminative, manifestant ainsi une hiérarchisation phrastique (Bonvini, 1992). Cet état des lieux, quoique partiel, met en exergue le débat théorique et conceptuel que soulève l'approche de ces constructions. Ces controverses mises à part, on retrouve chez la quasi-totalité des descripteurs les mêmes critères définitoires de base : une séquence de verbes - noyau d'une proposition unique - qui :

¹ On peut se référer aux théories dites substratiste et superstratiste. La première regroupe des auteurs défendant la thèse du substrat africain des créoles selon laquelle ces structures seraient dérivées des langues de l'Afrique de l'Ouest, suite à la déportation esclavagiste des Africains de cette région vers les îles (Manessy, 1995 ; Muysken et Veenstra, 1995, Lefebvre, 1991). Quant à la théorie superstratiste, elle stipule que ces structures sont tributaires d'une forme de français parlé à l'époque de la formation du créole au sein même de la colonie (G. Hazaël-Massieux, 1991 ; Wittman et Fournier, 1983).

- expriment un seul événement ;
- dépendent d'un seul et même sujet ;
- sont traversés par les mêmes modalités (mode, temps, aspect) qu'elles soient intégrées ou non ;
- partagent la même polarité (positive ou négative) ;
- ne sont ni coordonnés ni subordonnés.

Les exemples suivants tirés du créole haïtien (désormais CH), langue réputée avoir le plus de CVS parmi les créoles à base française, illustrent cette définition :

- (1) Yo **kouri** **ale** **chèche** dlo.
Elles *courent* *vont* *cherchent* de l'eau
Elles se précipitent d'aller chercher de l'eau.
- (2) Li **prese** **desann** **ale** **remèt** liv la.
Il *se presse* *descend* *va* *rend* le livre
Il se dépêche d'aller rendre le livre.
- (3) **Pran** liv la **pote** **ban** mwen.
Prends le livre *porte* *donne* moi
Apporte-moi le livre.

Les CVS du CH peuvent être employées pour exprimer, entre autres, les mouvements, les déplacements (1), (2) ; pour introduire le bénéficiaire d'une action à l'aide du verbe donner - *ba, ban, bay* - (3), ... Ainsi, là où en français par exemple on utilise un lexème verbal ou divers procédés pour rendre les différents schèmes du procès, le CH recourt aux CVS du fait notamment de la faible productivité dérivationnelle qui caractérise les lexèmes verbaux et du caractère polysémique ou générique de ces derniers.

La nature complexe du prédicat, souvent associée à la notion de proposition, semble caractériser avant tout la structure sérielle du point de vue grammatical. "*À côté des énoncés simples à une proposition (avec un seul verbe) et des énoncés complexes à plusieurs propositions (autant de propositions que de verbes), il faut poser un énoncé simple à une proposition où le prédicat est assumé par plusieurs verbes ou plutôt une base verbale composée de plusieurs lexèmes verbaux*" (Bole-Richard, 1978 : 24). On peut comprendre que la construction sérielle qui s'organise autour d'une combinaison de lexèmes verbaux a exactement les mêmes propriétés syntaxiques, les mêmes aptitudes transformationnelles - comme la nominalisation, relativisation, focalisation - qu'une phrase indépendante ayant pour prédicat une forme verbale simple. Par exemple, les constituants nominaux en (1) - '*yo*' et '*dlo*' respectivement sujet et objet - peuvent se focaliser de la même manière que dans une phrase simple à prédicat verbal simple :

- (1') a. Se [yo] ki kouri ale chache dlo.
Ce sont elles qui *courent* *vont* *cherchent* de l'eau
Ce sont elles qui se précipitent d'aller chercher de l'eau.
- b. Se [dlo] yo kouri ale chache.
C'est de l'eau elles courent vont cherchent
C'est de l'eau qu'elles se précipitent d'aller chercher.

Au niveau de la construction syntaxique, l'énoncé (1) - voir plus haut - diffère de (1') - ci-dessous. Dans le premier cas, l'ensemble de la base verbale constitue une seule et même proposition ; tandis que dans le second cas, il s'agit de deux propositions distinctes reliées par le morphème '*pou*' (pour), avec reprise du sujet ou non dans la deuxième.

(1'') [Yo kouri] pou [(yo) ale chache] dlo.
 Elles courent pour (elles) aller chercher de l'eau
 Elles se précipitent pour aller chercher de l'eau.

On peut constater qu'à cette différence syntaxique, s'ajoute une nuance sémantique. En (1), l'ensemble traduit le déroulement d'un événement/d'une action unique avec spécification du mode de réalisation. En (1''), 'pou' (pour) non seulement sépare deux événements mais aussi introduit celui qui représente la finalité de la première (elles courent/se précipitent dans le but de ...). De plus, les verbes de la structure sérielle appartiennent au même constituant prosodique. Ils sont prononcés d'un seul tenant sans pause interne, contrairement aux propositions non sérialisées où les verbes sont séparés par une pause intonative.

3. Cadre théorique et hypothèses

3.1. Structuration sémantique et organisation syntaxique

On assigne généralement à la construction sérielle une fonction de décomposition du procès qui consiste à conférer à chacun des signifiants qui y sont impliqués une expression particulière. D'après certains auteurs, il s'agit d'une structure cognitive vu qu'une action complexe est analysée par le locuteur comme une succession d'actions simples (Givón, 1991). En effet, la décomposition du procès se fait de façon analytique par des signifiants verbaux distincts et non de façon intégrative, contrairement au français avec des constructions comme '*aller voir* un spectacle'. En ce sens, « Ce qui caractérise la sérialisation est le fait que la juxtaposition linéaire des lexèmes verbaux perd son caractère énumératif qui autoriserait la pluralité des sujets pour assumer une fonction analytique : désigner les signifiés pris en compte dans la construction du constituant verbal. » (Manessy, 1995 : 171). Cet auteur défend l'idée selon laquelle les CVS sous-tendent un mode de structuration sémantique de type analytique [$P = Sél + Sé2 + Sén$] dont la syntaxe, au niveau plus explicite, n'est qu'une pure configuration de ce mode d'organisation particulière [$P = v1 + v2 + vn$].

Nous supposons alors que les locuteurs créolophones gèreront difficilement les CVS dans leur activité traductionnelle vu que les langues en question ont des configurations sémantiques distinctes qui transparaissent dans des organisations lexico-syntaxiques différentes². La question sous-jacente à cette hypothèse est de savoir si ces langues imposent des contraintes syntaxiques et/ou sémantiques dans le passage de L1 à L2 et vice versa.

3.2. Relation L1-L2 et transférabilité

La thèse du transfert dans le cadre de la comparaison des langues et des cultures a fortement été débattue au sein de l'analyse contrastive. Lado (1957) laissait entendre que des erreurs commises en L2 par des apprenants reflétaient l'influence de la L1 ; ou du moins, la structure grammaticale de la L1 tend à être *transférée* à la L2. Dans cette même veine, Odlin (1989), suite à de vastes recherches sur le transfert, explique que des facteurs de typologie linguistique et le degré de différence entre L1 et L2 sont déterminants dans la conduite de transfert. Par contre, Dulay et Burt (1983) ont nié de façon radicale toute possibilité de transfert entre la L1 et la L2. Leur conception dénommée « l'hypothèse L2 = L1 » est proche du LAD³ de Chomsky. Une autre conception, celle de Kellerman (1979), laisse entendre que

² Rappelons qu'il n'y a pas de rapport de coordination ni de subordination entre les lexèmes verbaux ; de plus, chacun d'entre eux peut être un verbe plein et autonome dans une construction autre que sérielle.

³ Language Acquisition Device

toutes les formes de transfert prévisibles ne se produisent pas. D'après lui, il y a des contraintes qui déterminent la *transférabilité* (« transferability ») et c'est le sujet et non la structure de la langue qui détermine ces contraintes en fonction d'une psychotypologie.

Ces trois conceptions conduisent également à s'interroger sur le traitement des CVS du créole au français par les locuteurs considérés comme ayant l'usage des deux langues. Les stratégies de traduction donneront-elles lieu à des effets translinguistiques⁴ récurrents du fait principalement de la particularité structurelle de la L1 ? Ce processus est-il automatique (selon la conception de Lado) ? Ou encore la gestion de ces structures varient-elles en fonction des locuteurs (selon la conception de Kellerman) ?

4. Résultats

Si l'on retrouve dans le parler des Haïtiens bilingues certaines spécificités à tel point que l'on parle d'un français haïtien (Pompilus, 1974), des structures assimilables aux CVS n'y ont pas été identifiées de façon formelle jusqu'ici, contrairement à certaines variétés du français d'Afrique notamment en Côte d'Ivoire (Hattiger, 1981 :209)⁵. Dans la gestion des CVS par les sujets lors du passage de la L1 à la L2, nous avons relevé très peu de cas où ils ont tendance à « transférer » cette structure en français. Cette tendance consisterait à juxtaposer en français un certain nombre de verbes de manière à marquer le signifié. Cette transférabilité s'est avérée être non automatique. Le passage L1-L2 semble plutôt résulter d'une action complexe liée à un ensemble de facteurs dont le niveau de compétence dans les deux langues, la créativité des sujets, les stratégies de traduction individuelles, etc. En effet, tous les sujets ont eu recours à différents procédés lexicaux et/ou grammaticaux ou, le cas échéant, des stratégies d'évitement⁶.

Les données ci-dessous font ressortir les occurrences des procédés de traduction employés par les sujets (Tableau 1) et le lien qui semble exister entre ces procédés et les types de CVS (Tableau 2). Nous distinguons les procédés lexicaux (première partie) des procédés syntaxiques (deuxième partie). Nous faisons apparaître la fréquence des diverses occurrences au moyen des symboles présentés ci-dessous. Rappelons que ces occurrences sont calculées sur une base de trente traductions pour trente sujets différents.

- Ø : aucune occurrence
- +
- ++
- +++
- ++++

Le cadre général qui justifie ces inventaires est de tenter de déterminer la part respective du sémantique – comment les locuteurs restituent le sens encodé dans la langue de départ – et du syntaxique – les langues en présence imposent-elles des contraintes syntaxiques sur les locuteurs – dans l'activité traductionnelle créole haïtien / français.

⁴ Selon Kellerman (1979), il s'agit de phénomènes d'emprunt, d'interférence, de transfert, d'évitement de transfert, de perte de langue.

⁵ Exemple : *Je vais m'en aller vite au pays avec mon enfant.*

⁶ Nous appelons stratégies d'évitement les cas où les locuteurs évitent la structure sérielle ou du moins passent à côté du contenu sémantique qu'elle véhicule.

4.1. Tableaux

		EXEMPLES
PROCEDES	OCCURRENCES	<i>Kouri pran devan pote ale vann</i> <i>Courent devancent portent vont vendre</i>
Dérivation verbale	++	4. S’empressaient d’ap-porter les marchandises qu’elles vont vendre
Adv. et loc. adverbiales	+++	5. Accouraient au devant des gens pour proposer leur marchandises
Prép. et loc. prépositives	++++	6. Couraient pour aller devant pour aller vendre
Paraphrases verbales	++	7. Couraient pour aller vendre leurs marchandises
Périphrases verbales	++	8. Font la course pour aller plus vite vendre leurs marchandises
Coordination	+	9. Se précipitaient et transportaient des marchandises
Subordination	++++	10. Devançaient les gens afin d’aller vendre leurs marchandises
Expansion	+	11. Se précipitaient, les devançant, et transportaient des marchandises qu’elles vont vendre
Evitement	+	12. Portant leurs marchandises

Tableau 1 - Paradigme des stratégies de traduction / Exemples extrait du corpus

TYPES DE CVS PROCEDES	Manière <i>Tonbe mache prese</i> (Se mettre à marcher vite)	But <i>Vire tèt gade</i> (Se tourner pour regarder)	Mouvement, déplacement <i>Kouri pran devan pote ale vann</i> (Hâter le pas pour aller vendre)	Accomplissement <i>Redi monte rive</i> (Parvenir à monter)
Dérivation verbale	Ø	Ø	++	Ø
Adv et loc. adverbiales	++++	+	+++	+
Prép. et loc. prépositives	+++	++++	++++	++++
Paraphrases verbales	++	+++	++	++
Périphrases verbales	+	+	++	+
Coordination	Ø	Ø	+	Ø
Subordination	Ø	Ø	++++	Ø
Expansion	Ø	Ø	+	+
Evitement	+	+	Ø	+

Tableau 2 - Occurrences des stratégies de traduction par type de CVS

4.2. Analyses

Les différents procédés lexicaux et grammaticaux inventoriés permettent de comprendre qu’à un CVS de type « X » ne correspond pas forcément une structure grammaticale spécifique de L2 de type « Y » pour que la relation sémantique soit équilibrée. Autrement dit, les langues en présence ne semblent pas imposer de contraintes syntaxiques sur les locuteurs.

Nous constatons un fort degré d’utilisation de prépositions ou de locutions prépositives pour l’ensemble des CVS. Elles sont employées comme éléments de segmentation pour éviter la création en chaîne de plusieurs séquences de verbes en français de telle sorte qu’il s’établisse entre les éléments ou les groupes fragmentés un rapport de subordination entre un verbe

principal et un ou de plusieurs subordonnées, contrairement aux CVS qui, rappelons-le, ne sont ni coordonnées, ni juxtaposées (*cf.* exemple 6).

Un nombre considérable d'adverbes et de locutions adverbiales sont aussi utilisés par la majorité des sujets comme stratégie de transposition lorsqu'il s'agissait surtout de traduire des mouvements de rapidité, différents aspects, ou encore certaines modalités de la relation prédicative exprimée au moyen des CVS (*cf.* exemple 5). L'aspect, catégorie par laquelle l'énonciateur conçoit le déroulement interne d'un procès, est marqué en créole haïtien au moyen de particules, marqueurs prédicatifs (MP) préposées au verbe. Toutefois, certains lexèmes verbaux, suivant le contexte de leur emploi, ont la particularité de souligner l'angle du déroulement du procès. C'est le cas des lexèmes verbaux '*tonbe*', '*resi*', '*pran*', présents respectivement dans les structures sérielles *tonbe mache prese*, *resi rive*, *pran gade* que nous avons soumises à nos sujets. En effet, '*tonbe*' et '*pran*', placés chacun au début de la structure sérielle, ont la particularité de présenter le procès dans sa phase initiale et sous un angle duratif. En transposant ces structures en français, les sujets ont conséquemment utilisé des périphrases verbales « se mettre à ... », « commencer à ... » parmi, bien entendu, d'autres procédés.

L'emploi des paraphrases verbales semble résulter d'un processus de simplification où, contrairement aux CVS, un seul lexème verbal est utilisé pour signifier les différents schèmes du procès. Cette simplification (Ferguson, 1977) semble être le processus par lequel le sujet ajuste son comportement linguistique aux exigences de la communication, l'ajustement étant conçu comme un choix restrictif parmi les ressources offertes par la grammaire du français (*cf.* exemple 7).

Les verbes formés par dérivation lexicale par préfixation correspondent le plus souvent aux CVS de type V1 focal + V2 déictique où V2 apporte à V1 une spécification directionnelle, indiquant le sens du mouvement par rapport à l'énonciateur. Dans le texte que nous leur avons soumis, on ne retrouve qu'un CVS de ce type-là (*pote ale*) qui est traduit tantôt par *apporter*, tantôt par *trans-porter* (*cf.* exemple 4).

Pour ce qui est des stratégies d'évitement, elles semblent liées à une difficulté à appréhender la structure sérielle et à rendre son équivalent sémantique en français L2. Dans certains cas, les locuteurs l'ignorent complètement, c'est ce que nous avons appelé « équivalence zéro » ; dans d'autres cas, ils réduisent la valeur sémantique du procès. Ces stratégies d'évitement pourraient être la conséquence d'une incompétence linguistique ou tout au plus pourraient être dues à une incapacité à s'adapter aux CVS du créole haïtien dans l'activité traductionnelle.

5. Conclusion

En considérant la variété des procédés de traduction employés par les locuteurs pour restituer le sémantisme des CVS du créole haïtien en français, on aurait tendance à confirmer la primauté du sémantique sur le syntaxique. Même quand l'activité de traduction peut occasionner des pertes sémantiques ou des détournements de sens dus à un ensemble de facteurs dont le niveau de compétence des sujets, les équivalences rendues en français L2 sont, dans l'ensemble, justes ou très rapprochées des CVS de départ et ne portent pas la trace d'une transférabilité de la structure de L1. Néanmoins, nous avons constaté une forte occurrence de la préposition 'pour' comme stratégie consistant à éviter les successions verbales en français. Aussi, avons-nous cherché à savoir ce qui se passe lorsque ces mêmes valeurs sémantiques comprises en français sont traduites en créole haïtien. Les données dont nous disposons révèlent *a priori* que l'emploi des CVS en créole haïtien ne semble pas être motivé par des contraintes grammaticales liées à une configuration sémantique particulière. Les équivalences rendues de L2 à L1 portent à croire qu'il s'agit là aussi d'une stratégie discursive qui varie d'un locuteur à l'autre.

Références

- BOLE-RICHARD, Rémi (1978). Problématique des séries verbales avec application au gen, *Afrique et langage*, 10 : 24-47.
- BONVINI, Emilio (1992). La construction sérielle : une structure cognitive ? *Linguistique africaine*, 8 : 7-53.
- DULAY, Heidi C. & BURT, Marina K. (1978). Some remarks on creativity in language acquisition, in *Second language acquisition research: issues and implication*, Ritchie, William C. (Ed). New York: Academy Press, 65-98.
- FERGUSON, Charles & DE BOSE, Charles (1977). Simplified registers, broken languages, and pidginization, in *Pidgin and creole linguistics*, Valdman, Albert (Ed). Bloomington : Indiana University Press.
- GIVÓN, Talmy (1991). Serial verbs and the mental reality of 'event': Grammatical vs. cognitive packaging, in *Approaches to grammaticalization: focus on theoretical and methodological issues*, Traugott, Elizabeth & Heine, Bernd (Eds). Amsterdam : John Benjamins, vol 1, 81-128.
- HAGÈGE, Claude (1975). Le problème linguistique des prépositions et la solution chinoise, *Société Linguistique de Paris LXXI*. Paris et Louvain : Peeters.
- HATTIGER, Jean-Louis (1981). *La morpho-syntaxe du groupe nominal dans un corpus de français populaire d'Abidjan*. Thèse de 3ème cycle de l'Université de Strasbourg.
- HAZAËL-MASSIEUX, Guy (1991). Genèse ou histoire de la modalité verbale en créole de la Guadeloupe, in *Modélisations en langues étrangères*, Russier, C., Stoffel, H., Véronique, D. (Eds). Aix-en-Provence : Publications de l'Université de Provence, 1996, 221-248.
- KELLERMAN, Eric (1979). La difficulté, une notion difficile, *Encrages*, n° spécial de linguistique appliquée : 16-21.
- LADO, Robert (1957). *Linguistics accross culture : Applied Linguistics for Language Teachers*, Ann Arbor, Michigan : University of Michigan.
- LEFEBVRE, Claire (1991). Take serial constructions in Fon, in *Serial verbs : grammatical, comparative and cognitive approaches*, Lefebvre, Claire (Ed). Amsterdam : John Benjamins Publication Company, vol 8, 37-79.
- MANESSY, Gabriel (1995). *Créoles, pidgins, variétés véhiculaires : Procès et Genèse*. Paris : CNRS Editions.
- MUYSKEN, Pieter & VENNSTRA, Tonjes (1995). Universalist approaches in *Pidgins and creoles : An introduction*, Arends, Jacques, Muysken, Peter & Smith, Norval (Eds). Amsterdam: John Benjamins, 122-134.
- ODLIN, Terence (1989). *Language transfer*. Cambridge : Cambridge University Press.
- POMPILUS, Pradel (1974). *La langue française en Haïti*. Port-au-Prince : Editions Fardin.
- VALDMAN, Albert (1978). *Le créole : structure, statut et origine*. Paris : Klincksieck.
- WITTMAN, Henry & FOURNIER, Robert (1983). Le créole, c'est du français, coudon! *Revue de l'Association Québécoise de Linguistique*, 3 : 187-202.

LES PROCÉDÉS DE MÉTAPHORISATION DANS LE DISCOURS MÉDICAL ARABE : ÉTUDE DE CAS

Tatiana EL-KHOURY

khourytatiana@yahoo.com

CRTT-ELISA, Université Lumière- Lyon2, 86, rue Pasteur
69365 Lyon

Abstract : This paper examines, through a large corpus of medical texts on organ transplants, the use of metaphors and their importance in written Arabic medical discourse. It focuses on the most typical metaphors of this field in order to establish a typology of the metaphorization process. It aims, thus, to explore the modalities for conceptualizing a scientific terminology in Arab societies.

1. Introduction

On s'attendrait, *a priori*, à ce que le discours scientifique soit neutre et sec, sans effets de style ni images métaphoriques. En réalité, il n'en est rien car, comme le constate à juste titre Rastier (1988 : 55), « l'histoire des sciences pullule de métaphores ». Et ce constat se vérifie clairement en médecine où la métaphore constitue un puissant outil de catégorisation, de dénomination et de vulgarisation. Ce dernier aspect (la vulgarisation) est primordial dans certains domaines médicaux où la métaphore permet de simplifier des notions complexes et dispose d'un fort potentiel imagé dans la représentation des concepts. En arabe, le discours de la transplantation présente un excellent cas d'illustration de cette emprise métaphorique où les termes imagés sont légion. On y retrouve des métaphores empruntées au domaine des finances, de la couture, de la propreté, etc.

Mais le principal domaine qui structure les images métaphoriques du discours de la greffe demeure, sans aucun doute, celui de l'agriculture. La pensée scientifique est à cet égard « bricoleuse » comme le note Schlanger (1971 : 20) ; elle va « emprunter ses éléments de construction là où elle les trouve, autour d'elle ». L'agriculture semble avoir, à cet égard, une place de choix dans l'univers de la vie courante. C'est cette métaphore agricole terminologique qui nous intéresse dans cet article à travers des exemples puisés dans un corpus de textes arabes portant sur la transplantation¹. Sur les 437 termes métaphoriques retenus suite à un dépouillement manuel², 202 se rapportent, en effet, au domaine de l'agriculture, soit près de 47% des termes métaphoriques du discours de la greffe. Nous allons nous pencher, plus précisément, sur les principaux procédés de métaphorisation, en tant qu'opération de conceptualisation ; l'objectif étant de mieux comprendre comment les Arabes conceptualisent et sémiotisent³ des notions médicales véhiculées par un discours de vulgarisation tel que celui de la greffe d'organes.

¹ Il s'agit du corpus électronique Transplant.1 constitué d'environ 600,000 mots. Ce corpus renferme des documents authentiques provenant de dix pays arabes (Arabie Saoudite, Koweït, Qatar, Emirats Arabes Unis, Liban, Syrie, Egypte, Maroc, Tunisie, Libye) et couvrant la période 1976-2005.

² Le dépouillement de Transplant.1 a permis de relever 4850 termes, tous types confondus (termes simples, termes complexes, signes brachygraphiques etc.).

³ Nous reprenons ici les termes de Pottier (2001) qui présente une synthèse des mécanismes supposés du monde mental dans lequel se réalise le passage des perceptions élaborées (la conceptualisation) vers les choix sémiologiques, c'est-à-dire les signes linguistiques (la sémiotisation).

Peu de publications arabes ont traité de la métaphore terminologique, encore moins dans le domaine de la médecine. Dans une approche interlinguistique, les études concernant la langue arabe ont notamment portées sur les difficultés de traduction et le degré de traductibilité des métaphores entre deux langues culturellement éloignées (surtout l'arabe et l'anglais)⁴. Tout en tenant compte de l'apport de telles études, nous nous proposons dans cet article de dresser une typologie des procédés de métaphorisation dans le discours arabe de la greffe dans une approche à la fois inter et intra linguistique car au-delà de la langue, il s'agit d'examiner la structuration d'une pensée. Pour ce faire, nous nous inscrivons dans le cadre des études cognitives qui mettent l'accent sur l'importance de la métaphore dans la structuration conceptuelle.

Les sciences cognitives ont établi un certain nombre de concepts pour expliquer les mécanismes de raisonnement, parmi lesquels la « métaphore conceptuelle » étudiée par Normand (1976), Lakoff & Johnson (1985) et appelée plus tard par Sinding (1993) « métaphore cognitive ». Pour les sémanticiens cognitivistes, les structures métaphoriques sont le reflet langagier d'un processus qui se produit non pas au niveau des mots mais de la pensée. C'est pourquoi, dans le processus de métaphorisation, l'opération de conceptualisation occupe une place de choix. Mais avant d'aborder les différents procédés de métaphorisation, qu'en est-il du rôle de la métaphore agricole dans le domaine médical et de la prégnance de l'agriculture (domaine-source où les concepts sont familiers) en transplantation (domaine-cible où les concepts sont inconnus) ?

2. La métaphore agricole en médecine et en transplantation

2.1. La métaphore agricole en médecine

La métaphore agricole est très prisée dans le domaine médical. De nombreux termes métaphoriques ou à composante métaphorique proviennent des domaines de la botanique, de la dendrologie (étude des arbres), de l'agriculture, etc. : *culture bactériologique*, *implantation dentaire*, *tronc cérébral*, *branche collatérale*, *rameaux nerveux*, *flore intestinale*...

Cette métaphore agricole est particulièrement présente dans le domaine de l'anatomie où elle remplit un rôle de description. En mobilisant notre faculté visuelle, la métaphore facilite la représentation mentale d'une réalité observée.

Cette expérience sensorielle s'avère également fondamentale pour la compréhension de réalités aux contours moins bien délimités. C'est le cas de la transplantation où il ne s'agit plus de décrire l'aspect ou la forme de tel ou tel organe mais de rendre compte d'un acte chirurgical divisé en deux phases principales, le prélèvement et la greffe, ainsi que de ses différents protagonistes (donneur, receveur, médecin greffeur etc.). L'outil métaphorique canalise alors plus que nos sens : notre faculté cognitive.

2.2. La métaphore agricole dans le domaine de la transplantation

Dans le domaine de la transplantation, de nombreux termes sont métaphoriques. Tout un pan est ainsi directement repris du vocabulaire agricole (*greffe*, *greffon*, *transplantation*, *implantation*...). L'émergence de ces métaphores a répondu à deux besoins essentiels : d'une part, il fallait nommer des réalités jusqu'alors inconnues et d'autre part il fallait que ces dénominations établissent des ponts entre les spécialistes et le public profane car la transplantation est un domaine à fort ancrage social. Mis à part les cas d'autogreffes où le receveur est le donneur lui-même, la transplantation demeure une pratique qui tire son

⁴ Voir à cet égard Davies & Bentahila (1989), Menacere (1992), Al-Harassi (2003) et ali (2006).

essence d'un attachement profond aux valeurs de solidarité, d'altruisme et d'entraide. Sans donneur, il n'y a point de receveur et donc pas de transplantation⁵.

C'est dans cette double optique à la fois de dénomination et de vulgarisation que la terminologie arabe de la transplantation a été créée. La plupart des publications médicales arabes relèvent, en fait, des ouvrages de vulgarisation⁶. Ainsi, la terminologie arabe de la greffe est essentiellement destinée au grand public dans le cadre des campagnes de sensibilisation au don d'organes. Cette terminologie recourt à la métaphore pour éveiller l'intérêt du public profane par la vivacité de l'image associée tout en mettant à sa portée une information complexe qui serait autrement rébarbative sinon inaccessible.

3. Les procédés de métaphorisation

Si la terminologie de la transplantation s'articule en français autour des verbes « greffer » et « transplanter », en arabe, il s'agit plutôt de « planter », de « semer », de « cueillir » et très rarement de « greffer » ou de « transplanter ». Pourtant, la créativité terminologique en arabe demeure essentiellement, à l'heure actuelle, une adaptation des termes arabes lors de leur traduction du français ou de l'anglais. Mais ce transfert linguistique ne se cantonne pas à des traductions littérales des unités terminologiques, notamment celles formées par métaphore, mais dépasse le calque pour s'imprégner de la culture arabe et s'adapter au système référentiel de la langue d'accueil. Les exemples métaphoriques puisés dans notre corpus, Transplant.1, étayent cette idée et font ressortir quatre procédés de métaphorisation dont deux sont interlinguistiques (nous proposons de les désigner par « métaphorisation par calque » et « métaphorisation par modulation ») et deux intralinguistiques (nous les appellerons « métaphorisation par extension » et « métaphorisation par énantiosémie »). Dans ce qui suit, nous tentons d'explicitier ces quatre procédés tout en soulignant l'importance de chacun d'eux dans la création terminologique métaphorique en arabe médical.

3.1. *La métaphorisation par calque*

Longtemps sous influence britannique et française, le monde arabe a vu sa langue scientifique fortement imprégnée de l'anglais⁷ (notamment dans les pays du Golfe et en Egypte) et du français (notamment dans les pays du Maghreb et au Liban) à travers des structures calquées et des emprunts.

La métaphorisation par calque consiste à « importer » une image métaphorique qui frappe l'esprit du locuteur arabophone parce qu'elle correspond à une image culturelle parlante. Dans ce type de métaphorisation, le terme de « calque » peut avoir deux implications distinctes. La première renvoie au transfert de la faculté métaphorique du français ou de l'anglais vers l'arabe, en utilisant des images qui sont censées effectuer des fonctions analogues au niveau cognitif dans les deux langues. Cette première implication correspond à

⁵ Nous excluons les rares cas de xénotransplantation où le donneur est d'une espèce différente (dans le cas des greffes humaines, il s'agit de prélever un organe sur un animal notamment le porc ou le babouin) ainsi que l'implantation d'organes artificiels délaissés au profit d'une nouvelle technique encore à l'essai qui consiste à cultiver des cellules (par exemple des cellules pancréatiques ou de la peau) et bientôt des organes entiers pour les greffer sur le patient.

⁶ La Syrie est le seul pays arabe à avoir entièrement arabisé son enseignement universitaire ce qui explique la rareté des publications spécialisées en la matière.

⁷ Il ne fait aucun doute que l'essentiel de la recherche scientifique est effectué aujourd'hui aux Etats-Unis et il incombe à toute personne désireuse d'approfondir ses recherches dans un domaine scientifique de recourir à une bibliographie en anglais. Ce phénomène est exacerbé dans le monde arabe où la grande majorité des facultés de médecine dispensent leurs enseignements en anglais.

la démarche consistant à créer des termes nouveaux répondant à des besoins pratiques, des métaphores fonctionnelles en quelque sorte.

La seconde implication concerne l'effet de cette image dans la langue-culture cible (ici l'arabe) et les capacités cognitives de méta-représentation dans les deux langues considérées (le français et l'arabe ou bien l'anglais et l'arabe). Cette deuxième implication correspond à la démarche consistant à reconstituer le cheminement mental des utilisateurs de la langue afin de retrouver les modalités et le processus selon lesquels les termes en usage ont été forgés et diffusés.

Ces deux démarches sont complémentaires et permettent d'appréhender la complexité de la terminologie de la transplantation et du processus de métaphorisation.

De façon plus concrète, la métaphore est *transposée* sémantiquement du français ou de l'anglais par le biais du calque. Il y a *transposition* quand le sémème⁸ tout entier passe de la langue source à la langue cible. Ainsi, le terme « greffe » est parfois rendu, par le biais du calque, par l'équivalent arabe /tat^cīm/⁹. Dans le processus de métaphorisation par calque, on passe d'un domaine cible dans une langue étrangère (la médecine en français) à un domaine cible dans une langue d'accueil (la médecine en arabe) en passant par les domaines sources des deux langues concernées (l'agriculture en français et en arabe). La circularité du processus se fait de la médecine en français vers l'agriculture en français (intra-langue A)¹⁰. Ensuite de l'agriculture en français vers l'agriculture en arabe (inter-langue). Enfin de l'agriculture en arabe vers la médecine en arabe (intra-langue B).

Le passage en intra-langue d'un domaine à un autre se fait par la mise en valeur d'un ou plusieurs sèmes particuliers et « la mise entre parenthèse » (Le Guern, 1973 : 15) d'autres sèmes constitutifs du lexème employé. C'est ce que Assal (1994 : 236) appelle le « mécanisme d'abstraction sémique » dans lequel se réalise un tri sémantique qui occulte les sèmes qui ne servent pas l'image métaphorique et met en valeur ceux qui, par leur adéquation, établissent l'analogie et légitiment la métaphore.

Pour illustrer ce phénomène de mise en valeur sémique, prenons tout d'abord le terme « greffe » dans les deux domaines source et cible. En médecine, la greffe est : « L'opération qui consiste à **insérer**¹¹ une portion de l'organisme d'un individu (donneur) sur une autre partie du corps (autogreffe) ou sur un autre individu (receveur) (allogreffe) » (*Petit Robert*).

Dans son domaine d'origine, l'agriculture, la greffe signifie : « Opération par laquelle on **implante** un greffon (greffon : Partie d'un végétal dont on veut obtenir de nouveaux spécimens et qu'on greffe sur un autre végétal (porte-greffe)) » (*Petit Robert*).

Pour suivre le chemin du calque métaphorique, il faudrait donc retrouver les sèmes mis en valeur afin de les réinvestir lors du transfert linguistique à la langue cible. Dans le cas de la « greffe », il s'agit notamment de ce mouvement « d'insérer, d'implanter » qu'on retrouve dans la définition arabe de /tat^cīm/ dans le domaine agricole sous la forme du verbe « attacher » : « Greffer une branche : c'est lui attacher une branche, un bourgeon ou tout

⁸ Nous empruntons à Pottier (1974) le terme de « sémèmes » qu'il faudrait comprendre comme des « paquets de sèmes ». Les linguistes français parlent souvent, avec Pottier et Greimas, de « sème », celui-ci n'étant ni un signifiant ni un signifié mais une unité sémantique minimale non décomposable (Ducrot & Schaeffer, 1995 : 533-534).

⁹ Les symboles de transcription utilisés dans cet article sont ceux de l'*Encyclopédie de l'Islam*. La transcription de certains caractères de l'arabe n'étant pas possible avec nos logiciels de traitement de texte, les phonèmes emphatiques sont mis en gras : **d**, **k** et **t**.

¹⁰ C'est le chemin inverse de la création métaphorique originale qui est partie, dans ce cas précis, de l'agriculture en français vers la médecine en français.

¹¹ C'est nous qui mettons en gras.

autre partie d'une espèce différente afin qu'elle devienne de cette même espèce » (Dictionnaire *Ar-Ra'id*).

C'est ce sème de fixation d'un élément étranger à un autre, cette relation binaire de deux éléments qui s'incorporent, manifeste dans les définitions du terme « greffe » en arabe comme en français qui établit l'analogie métaphorique et permet la transposition par le calque. Ainsi en arabe médical, on obtient /tat^cīm/ « la greffe » mais également son dérivé /tu^cm/ « le greffon ».

Transcription du terme arabe	Glose littérale ¹²	Equivalent français
tat ^c īm	Grefe	Grefe
tu ^c m	Grefon	Grefon
tu ^c ūm	Greffons	Greffons

Toutefois, ce terme et ses dérivés sont rarement utilisés dans le domaine médical arabe de la transplantation et s'appliquent de façon quasi-exclusive à la greffe de la peau. Les arabophones leur préfèrent de loin un autre terme et ses dérivés; il s'agit de /zar^c/ « action de planter » obtenu par modulation.

3.2. La métaphorisation par modulation

Les études comparatives entre systèmes linguistiques démontrent que les termes considérés comme des équivalents ne renvoient pas systématiquement à des concepts identiques. Comme le rappelle Oliveira (2005 : 284-285) « l'équivalence terminologique ne présuppose pas forcément l'équivalence notionnelle ». D'un point de vue traductologique, la modulation est le procédé privilégié pour l'étude des différences notionnelles entre langues et des variations conceptuelles entre cultures mises en contact par le biais de la traduction.

La « modulation » est le terme proposé par Vinay et Darbelnet (1958 : 51) pour désigner un certain nombre de variations qui deviennent nécessaires quand le passage de la langue source à la langue cible ne peut se faire directement. « Moduler » signifie adapter la traduction à des cadres thématiques et culturels particuliers. En général, elle consiste à introduire dans la langue cible une variation dans le point de vue afin que le résultat de la traduction soit conforme aux usages des récepteurs finaux du message. Vinay et Darbelnet (1958 : 89) précisent que la modulation sert à agir sur les « catégories de la pensée » et distinguent onze types de modulation¹³. Parmi ces types nous nous intéressons à la modulation substitutive dans laquelle une partie est traduite par une autre.

Pour expliciter notre propos, nous prenons l'exemple de /zar^c/, le nom verbal du verbe /zara^ca/ qui signifie étymologiquement « semer » mais qui est utilisé en arabe moderne au sens de « planter ». À l'instar de /tat^cīm/ « greffe », /zar^c/ « action de planter », est un acte agricole. Mais chez les Arabes, /zar^c/ est plus prototypique¹⁴ que /tat^cīm/ car ce terme est primordial dans la catégorie des actes agricoles.

De plus, les deux termes sont polysémiques dans le discours médical ; /zar^c/ signifie à la fois « greffe » et « culture » et /tat^cīm/ signifie « greffe » et « vaccination ». Toutefois, de par son usage, /tat^cīm/ est perçu davantage comme l'équivalent de « vaccination » (le sens le plus courant) alors que /zar^c/ est perçu avant tout comme l'équivalent de « greffe ». Ainsi nous retrouvons une prédominance du terme /zar^c/ dans la grande majorité des corpus arabes

¹² Les termes arabes sont glosés en français par nos soins.

¹³ Pour une présentation complète des onze types de modulation, voir Vinay et Darbelnet (1958 : 89-90).

¹⁴ Dans la version standard du prototype, « le prototype est conçu comme étant le meilleur exemplaire communément associé à une catégorie » (Kleiber, 1990 : 49).

excepté dans le corpus syrien où il est concurrencé par un autre terme synonyme /ghars/ qui signifie également « planter ».

Transcription du terme arabe	Glose littérale	Equivalent français
zara ^c a	Semer (sens étymologique) Planter (sens moderne)	Greffer
zar ^c	Action de planter (=plantation)	Greffer

Les spécialistes arabes et français de la transplantation interprètent donc différemment un même concept en raison de leur bagage culturel. Le calque devient alors moins fréquent dans des sociétés de cultures éloignées en raison d'une absence de parallélismes de pensée et de références métaphoriques semblables et c'est la modulation qui prend la relève.

Une fois transférée par modulation en arabe, la métaphore peut étendre les catégories existantes en ajoutant de nouveaux termes qui n'étaient pas métaphoriques dans la langue source. Commence alors un travail intralinguistique au sein même de l'arabe pour élargir, par proximité thématique, le champ sémantique d'un terme productif et bien implanté. Ce type de métaphorisation semble se faire par « extension » comme nous allons tenter de le montrer ci-après.

3.3. La métaphorisation par extension

La métaphorisation par extension consiste à étendre une image issue d'un champ sémantique exploité partiellement dans d'autres langues (en l'occurrence le français ou l'anglais) vers d'autres termes de ce même champ. Cette opération permet de greffer sur l'information référentielle initiale des images annexes facilement saisissables. Ainsi, par exemple, les spécialistes arabes de la greffe vont étendre l'image de /zar^c/ « action de planter » à d'autres termes du domaine comme le montre le tableau suivant:

Transcription du terme arabe	Glose littérale	Equivalent français
al-katf	La cueillette	Le prélèvement
al-katūf	La cueillette	Le prélèvement
al-kulya al-maktūfa	Le rein cueilli	Le rein prélevé

La métaphore terminologique est le résultat d'une construction chargée de valeurs culturelles. En d'autres termes, lorsque des significations et des associations ont cours dans un cadre culturel, elles risquent de se refléter dans des images métaphoriques propres à une langue donnée. Autant dire que la métaphore s'appuie sur des données culturelles familières pour pouvoir investir de nouveaux champs de connaissances. Ainsi, lorsque le spécialiste est conscient de l'implantation d'une image métaphorique dans une culture donnée, il est possible de jouer sur celle-ci pour élargir la portée d'un concept initial. Dans le domaine de la transplantation arabe, par exemple, les termes métaphoriques exprimant le prélèvement vont également puiser dans le domaine de l'agriculture. Ils font appel à l'image évocatrice de la « cueillette » qui complète le tableau agricole. On « cueille » un organe pour le « planter » ailleurs. N'est-ce pas le propre du mouvement double de la transplantation ?

Mais il arrive que la métaphorisation intralinguistique ait recours à des phénomènes connus traditionnellement en arabe classique, tels que l'énantiosémie. Celle-ci se retrouve ainsi parmi les principaux procédés de métaphorisation dans le discours arabe de la greffe.

3.4. La métaphorisation par énantiosémie

La métaphorisation par énantiosémie consiste à utiliser un terme métaphorique dans deux sens directement opposés l'un à l'autre. L'énantiosémie est un phénomène linguistique qu'on observe dans toutes les langues sémitiques et en vertu duquel un mot possède deux sens contraires. Ces mots ayant deux sens antithétiques sont désignés par les philologues arabes par le terme /addad/ (s. /did/). L'opinion qui prévaut en linguistique générale est que l'énantiosémie ne peut être expliquée en fonction d'un principe unique. Selon l'*Encyclopédie de l'Islam* (1991 : 191), « les mots ont depuis l'origine une signification établie ; dans le cas

de chaque /did/, il faut donc considérer l'un des deux sens comme primitif et l'autre comme adventice ».

Dans le domaine de la transplantation, une forme en particulier peut avoir une signification opposée à celle qu'elle a usuellement : il s'agit de /zāri^c/. Ce mot possède comme sens primitif celui impliqué par son schème¹⁵ /fā^cil/ soit l'agent de l'action, en l'occurrence « celui qui plante quelque chose, le planteur ». Mais on constate dans les textes de notre corpus un sens concurrent et antonyme de ce même mot qui est « celui qui subit l'action de planter » sans avoir toutefois le schème du nom de patient /maf^cul/. Ce sens adventice (« receveur » pour /zāri^c/) pourrait s'expliquer par une focalisation et un changement de point de vue : dans le premier cas (sens primitif), le point de vue met en relief le médecin greffeur comme agent de l'action de planter et par conséquent le patient est vu comme /mazrū^c/ « celui qui subit l'action de planter » (sous-entendu par le médecin). En revanche, dans le deuxième cas (sens adventice), l'agent de l'action de planter est perçu comme le patient lui-même, présenté comme actif dans son opération : il est par conséquent le « planteur » /az-zāri^c/.

Transcription du terme arabe	Glose littérale	Equivalent français
az-zāri ^c	Le planteur	Le receveur
az-zāri ^c ūn	Les planteurs	Les receveurs
az-zāri ^c a	La planteuse	La receveuse
az-zāri ^c āt	Les planteuses	Les receveuses

4. Conclusion

Ces différents procédés de métaphorisation aident à mieux comprendre comment les usages métaphoriques de la greffe sont conceptualisés et sémiotisés en arabe médical. Il en ressort que la création métaphorique dans le discours de la transplantation relève d'un double processus interlinguistique et intralinguistique et tend davantage à l'innovation culturellement acceptable qu'au calque. La terminologie arabe de la transplantation, en tant que vaste métaphore filée du domaine agricole, semble être un moyen privilégié pour capitaliser sur l'importance de l'agriculture, principale source de revenu de plus de 30% des populations arabes¹⁶, permettant ainsi d'investir, grâce à la prégnance des images agricoles, un autre domaine tout aussi vital, la médecine.

Références bibliographiques

- AL-HARASSI Abdulla (2003). Ideological Conceptual Metaphors in Arabic-English Translation of Political Discourse. *International Journal of Arabic-English Studies*, 4 : 85-105.
- ALI Abdul Sahib Mehdi (2006). On The Translation of Metaphor: Notions and Pedagogical Implications. *International Journal of Arabic-English Studies*, 7 : 121-136.
- ASSAL Allal (1994). La métaphorisation terminologique. *Terminologie et traduction*, 2 : 235-242.
- DAVIES Eirlys E. & BENTAHILA Abdelai (1989). Familiar and Less Familiar Metaphors: An Analysis of Interpretations in Two Languages. *Language and Communication*, 9 (1) : 49-68.
- DEBOVE-REY Josette & REY Alain (2001). *Le Nouveau Petit Robert*. Paris : Dictionnaires le Robert.

¹⁵ Le schème est le « moule » syllabique qui sert à la formation des mots en arabe. Selon le schème combiné avec la racine du verbe, nous pouvons obtenir le nom verbal, le nom d'agent (ou participe actif), le nom de patient (ou participe passif), le nom de lieu, le nom d'instrument, le nom d'abondance, le nom de métier etc. Pour plus de détails, voir Guidère (2001).

¹⁶ Selon le rapport pour l'année 2004 de l'Organisation arabe pour le développement agricole.

- DUCROT Oswald & SCHAEFFER Jean-Marie (1995). *Nouveau dictionnaire encyclopédique des sciences du langage*. Paris : Edition du Seuil.
- GIBB Hamilton Alexander Rosskeen *et al.* (1991). *Encyclopédie de l'Islam*. Leiden : Brill.
- GUIDÈRE, Mathieu (2001). *Grammaticalement correct ! Grammaire alphabétique de l'arabe*. Paris : Ellipses.
- KLEIBER Georges (1990). *La sémantique du prototype*. Paris : Presses Universitaires de France.
- LAKOFF George, JOHNSON Mark (1985). *Les métaphores dans la vie quotidienne*. Paris : Editions de Minuit.
- LE GUERN Michel (1973). *La sémantique de la métaphore et de la métonymie*. Paris : Larousse.
- MAS'UD Jibran (2003). *Ar-Ra'id*. Beyrouth : Dār al-ʿilm lil-malāyīn.
- MENACERE Mohammed (1992). Arabic Metaphor and Idiom in Translation. *Meta*, 37 (3) : 567-572.
- NORMAND Claudine (1976). *Métaphore et concept*. Bruxelles : Ed. Complexe.
- OLIVEIRA, Isabelle (2005). *Nature et fonctions de la métaphore dans la terminologie médicale. Etude comparée du français et portugais*. Thèse. Lyon : Université Lumière Lyon II.
- POTTIER Bernard (1974). *Linguistique générale. Théorie et description*. Paris : Klincksieck.
- POTTIER Bernard (2001). *Représentations mentales et catégorisations linguistiques*. Leuven : Peeters.
- RASTIER, François (1988). Paradigmes cognitifs et linguistique universelle. *Intellectica*, 6 : 43-74.
- SCHLANGER, Judith (1971). *Les Métaphores de l'organisme*. Paris : Librairie Philosophique J. Vrin.
- SINDING Christiane (1993). Les métaphores en biologie – analogies ou outils de pensée ? *Intellectica*, 16 : 85-99
- VINAY Jean-Pierre & DARBELNET Jean (1958). *Stylistique comparée du français et de l'anglais*. Paris : Didier.
- WEIL Gustav (1991). Addād. *Encyclopédie de l'Islam* : 191-192.

LE CONCEPT DE *PHRASE CANONIQUE* EN LINGUISTIQUE FRANÇAISE

Antoine GAUTIER

antoine.gautier@paris4.sorbonne.fr

EA 4089 Sens, Texte, Histoire

Université Paris Sorbonne

75005 Paris

Abstract : The concept of ‘canonical sentence’ is quite common in French linguistics and grammar. Its current definition remains fuzzy and the how and why of its use also remains unclear. The philosophy of science has much to tell about this kind of theoretical artefact, and in particular it seems able to dissociate several functions of the canonical sentence, such as a metalinguistic use.

1. Introduction

Très peu de grammaires du français contemporain font l'économie d'une allusion à la *phrase canonique* (désormais PC) que ce soit explicitement ou par l'entremise d'une formulation approchante comme *phrase type*. Pour peu que l'on confronte les différents usages d'un ouvrage à l'autre, on constate que le concept est instable et mal délimité, et qu'en outre il prend place dans une procédure descriptive incertaine car trop intuitive. Aussi semble-t-il essentiel de proposer une définition précise du concept et de sa place dans le métadiscours linguistique, en voyant en particulier à quoi rattacher cette « canonicité » dont l'acceptation spontanée est suspecte. Cette démarche est susceptible de permettre la différenciation des outils épistémologiques que, selon nous, la PC recouvre indistinctement.

2. La phrase canonique dans les grammaires de référence

À la suite de C. Germain et H. Séguin (1998), on distingue parmi les grammaires explicites, c'est-à-dire les tentatives de notation de la compétence, les grammaires dites **pédagogiques**, inscrites dans le processus d'enseignement ou d'acquisition d'une langue, et les grammaires **linguistiques**, qui visent à décrire ou à simuler au mieux cette langue. Pour l'essentiel, c'est parmi ces dernières, dans les grammaires **descriptives**, qu'il est recouru à la notion de PC, ce qui invite à voir une corrélation entre les objectifs spécifiques de ce type d'ouvrage et l'objet qui nous occupe ici.

Trois références ont nourri notre enquête : la *Grammaire méthodique du français* (Riegel *et al.*, 1994), la *Grammaire de la phrase française* (Le Goffic, 1993) et la *Grammaire du français* (Denis & Sancier-Château, 1994). Un bref parcours révèle que la PC est systématiquement convoquée pour définir la **phrase** elle-même¹ – autre insaisissable du métalangage. En outre, le fait que chaque grammaire l'emploie sous des *noms* et/ou des *formes* différentes n'est pas étranger à l'instabilité de la notion.

2.1. Appellations et représentations

Riegel *et al.* évoquent successivement une « phrase de base », une « phrase prototypique » et la « phrase canonique » ; ils proposent en outre une représentation en métalangage

¹ À titre d'exemple : « Il est avantageux de se doter au départ d'une structure (ou forme) canonique de la phrase. [...] C'est par rapport à ce schéma de référence (ou modèle opératoire) que sont décrites toutes les phrases observables. » (Riegel *et al.*, 1994 : 109).

syntactique : (CC) — *Sujet* — (CC) — *Verbe* — *Complément(s) / Attribut* — (CC). La *Grammaire du français* fait état d'une « phrase archétypale » apparentée à la phrase verbale simple. Enfin, la *Grammaire de la phrase française* produit une occurrence de « phrase type » : *Marie chante*.

Bien que leur fonction commune autorise à voir dans ces entités les avatars d'un même concept, la diversité des **appellations** n'est pas dénuée de répercussions théoriques : par exemple, parler d'une « phrase de base » implique que d'autres formes dérivent d'une même structure, comme l'interrogation et la négation dériveraient de l'assertion positive (la notion de transformation n'est alors plus très loin). D'autre part, avec l'étiquette de « phrase prototypique », le problème semble désert l'objectivité des faits pour une forme de subjectivité propre au jugement d'appartenance catégorielle (voir *infra*).

Problème crucial posé à toute science dotée d'ambitions théoriques, les **représentations** ont elles aussi un impact sur le concept qu'elles dénotent et sur son articulation au reste du modèle : ainsi Riegel *et al.* ont-ils recours ci-dessus au métalangage formel pour proposer une visualisation de la PC. Ils produisent en outre une définition en extension de l'objet en listant ses caractéristiques². Enfin, le choix de Le Goffic est de représenter la PC par l'entremise d'une occurrence de phrase, ce qui ressortit d'une définition en extension. Naturellement, ces choix théoriques sont autant d'accès différents à un seul et même objet : la phrase canonique. La **représentation métalinguistique** se montre en tant que structure syntaxique abstraite, tandis que la **description métalinguistique** dit cette même structure et que l'**exemple** la réalise effectivement, passant du type à l'occurrence. Y a-t-il parmi ces options un choix moins pertinent ? Force est d'admettre, si l'on suit Riegel *et al.* qui veulent voir dans la PC un **modèle**, que la description remplit mal un tel office, puisqu'elle énonce les propriétés de son objet sans montrer cet objet lui-même.

2.2. *Structure de la phrase canonique*

Si l'on laisse de côté ces divergences, il subsiste encore d'importants écarts entre les phrases canoniques des grammaires étudiées. À titre d'exemple, entre les phrases « *Marie chante*. » et « *Marie chante une chanson*. », peut-on déclarer que l'une est « plus » canonique que l'autre ?

On peut tenir pour évident que la PC est avant tout un modèle syntaxique et non pas sémantique. C'est pourquoi il importe peu que l'on rencontre « *Marie chante*. » ou « *Marie mange*. », qui semblent canoniques au même degré. À l'inverse, une phrase plus atypique, comme « *Marie abat un sapin*. », sera moins aisément éligible comme PC du fait que son sens dénotatif fait obstacle à une lecture canonique³. Mais le plus important est que, d'une grammaire à l'autre, les critères constitutifs du « canon » divergent même d'un point de vue syntaxique : la *Grammaire de la phrase française* retient une phrase comportant un Verbe et un Nom propre sujet (« *Marie chante*. »), tandis que la *Grammaire méthodique du français* y ajoute, dans la représentation citée plus haut, un complément d'objet (ou un attribut) et d'éventuels circonstants.

En fait, on gagne ici à distinguer deux notions qui se côtoient couramment dans les grammaires de référence, à savoir la *phrase canonique* et la *phrase minimale*. La seconde constitue non pas un modèle de la catégorie des phrases mais plutôt une forme de limite structurelle de cette catégorie, en deçà de laquelle, dans une conception très normative, on ne

² « Phrase assertive, simple [...] et neutre. » (Riegel *et al.*, 1994).

³ À cet égard, la sémantique ne semble pas complètement étrangère à notre propos. Pour autant, la PC ne saurait être qu'une séquence décontextualisée et non un « énoncé » dans la mesure où elle illustre des faits de *langue* et non de *parole*.

peut plus parler de « phrases ». À l'opposé, la structure de Riegel *et al.* est plus complexe et permet de décrire contrastivement d'autres types de phrase de la grammaire française ou d'autres langues. Autrement dit, la *phrase minimale* occupe une frontière de la catégorie et la PC en occupe le centre.

Concernant la forme syntaxique de la PC, il est éloquent que le modèle typologique de Greenberg (1966) comporte non pas deux mais trois constituants (SVO). Cela invite à voir dans le schéma proposé par Riegel *et al.* une structure « plus » canonique que celle de la phrase à deux termes. En combinant les critères définitoires des grammaire évoquées ici et ceux d'autres textes (dont Willems, 2001), on est en mesure de proposer une synthèse des propriétés que semble fixer la phrase canonique : il s'agirait donc d'une *Phrase assertive délimitée graphiquement ou prosodiquement, neutre, simple et d'ordre SVO*.

3. Définir la canonicité ?

Les différences terminologiques étant à peu près aplanies, il reste à déterminer précisément pourquoi la PC a la structure mentionnée, c'est-à-dire celle d'une phrase assertive simple, et non pas celle d'une phrase interro-négative clivée, par exemple. Certains auteurs avancent le droit à l'axiome et ne justifient pas leur choix ; d'autres, tels Riegel *et al.*, proposent quelques éléments d'explication :

(i) ils avancent tout d'abord que cette forme est la plus **immédiate**. Puisqu'il semble être question d'une chronologie – dont la nature reste à préciser – il faut voir dans la PC une structure primordiale dont procèdent les autres types de phrase (interrogatif, exclamatif, clivées, etc.) ; dans les premières versions de la grammaire générative, la notion de *phrase noyau*, obtenue par un nombre minime de transformations, s'approchait nettement de cette conception. Selon la nature de la chronologie que sous-entend l'*immédiateté* invoquée, on fera tantôt de la PC une structure de base de la grammaire intériorisée, de la compétence, tantôt un outil descriptif de linguiste. Dans le premier cas, la PC est réputée prégnante dans les processus cognitifs et constitue une réalité psychique ; dans l'autre, son antécédence est liée à une formalisation particulière et elle reste un artéfact théorique. Quoi qu'il en soit, cette hypothèse est difficilement falsifiable ou vérifiable, mais elle est surtout très engagée théoriquement.

(ii) Une deuxième voie d'explication se fonde sur la fréquence relative de la structure canonique. L'idée générale est que l'on voit se réaliser plus souvent, dans le discours, des phrases assertives simples que, par exemple, des interrogatives ou des phrases averbales. Si cela se vérifie pour les pages du *Monde*, comme l'a montré Ulland (2002), c'est plus difficile à concevoir pour la langue littéraire ; quant à l'oral, Lambrecht (2007) a montré récemment une propension des locuteurs à **ne pas** produire de structures canoniques.

(iii) La troisième option envisagée par Riegel *et al.* est de considérer la structure de la phrase canonique comme la plus **représentative** de la catégorie des phrases. Le problème se déplace alors du plan des faits linguistiques à celui du métalangage. Dans les deux premiers cas, les explications se fondaient sur la configuration supposée de la grammaire intériorisée ou sur l'observation des productions de discours. Ici, en l'espèce, c'est le rapport d'une représentation mentale à une catégorie de métalangage qui entre en ligne de compte et non le système de la langue lui-même.

4. Un modèle de phrase ?

On convoque souvent la PC comme un *modèle* de la phrase ; or, un tel concept ne semble pas devoir être pris légèrement compte tenu de ses répercussions théoriques. En premier lieu, l'acception doxique du « modèle » – « chose ou personne servant de référence à l'imitation »

(TLFi) – ne correspond pas même de loin à la fonction de la PC dans les grammaires de référence⁴. Une définition épistémologique semble plus rigoureuse :

« *Système représentant les structures essentielles d'une réalité et capable à son niveau d'en expliquer ou d'en reproduire le fonctionnement.* » (TLFi).

Mais alors survient un autre problème, relatif à la possibilité de « reproduire le fonctionnement » de l'objet. Deux conceptions du « modèle » se trouvent ici opposées : une version *faible* visant la seule représentation statique de certaines propriétés structurales, et une version *forte* qui y ajoute la dimension du fonctionnement. En linguistique, le passage d'une acception à l'autre correspond à la transition d'un système **descriptif**, qui représente des phrases, à un système **génératif**, capable de prédire la forme d'autres objets plus ou moins complexes, tout en opérant la discrimination des phrases bien ou mal formées. Les théories chomskyennes, qui font usage de modèles « forts », en sont une illustration, mais on peut voir de cela une version simplifiée dans la représentation proposée par Riegel *et al.*

Qu'en est-il des autres formes de PC ? Comme on l'a dit plus haut, la définition en intension, tout d'abord, est incapable de remplir un rôle de modèle – puisqu'elle ne *représente* rien. Quant à l'occurrence de PC « *Marie chante une chanson.* », il s'avère qu'on ne peut fabriquer d'autres phrases à partir d'elle (et encore moins en prédire l'acceptabilité) sans possession préalable de la grammaire complète de la langue.

En fait, la vocation de cette occurrence n'est pas de démontrer mais de montrer un certain nombre de propriétés positives : la phrase « *Marie chante une chanson.* » n'a pour seul rôle que d'incarner un exemplaire de phrase assertive neutre d'ordre SVO. Mais elle l'incarne en déployant autour d'elle un réseau de ressemblances et d'oppositions qui relie toutes les autres structures de phrase possibles de la langue et qui constitue la catégorie des *phrases*.

Suivant les travaux de Rosch (1973), on est fondé à voir dans cette occurrence de phrase un **prototype** de sa catégorie :

« *Un prototype est une instance typique d'une catégorie [...] les autres éléments sont assimilés à cette catégorie sur la base de leur ressemblance perçue avec le prototype.* » Langacker (*apud Kleiber, 1990*).

Dans le cadre d'une grammaire de référence, qui vise à décrire et expliquer le code, un tel outil a pour fonction de fonder le métalangage non pas sur des propriétés réputées inhérentes aux objets, c'est-à-dire sur le système des conditions nécessaires et suffisantes, mais sur des représentations mentales partagées. C'est du reste cette même phrase prototypique qui est utilisée dans la typologie de Greenberg (1966) : pour comparer et classer les langues, on confronte des occurrences de phrases éligibles comme instances typiques de la catégorie des *phrases* de chaque langue.

5. Conclusion

Un bref parcours à travers quelques grammaires révèle bien une tendance commune à utiliser un **concept régulateur** pour aborder la question de la phrase. Ce concept connaît des appellations et des représentations variées : la *Grammaire de la phrase française* propose ainsi un outil plus proche de la phrase minimale, tandis que les autres font usage, sans distinction explicite, de modélisations (faibles ou fortes) et de prototypes.

Toutes ces solutions visent à contourner le caractère insaisissable de la notion de « phrase ». Comme l'a montré récemment Kleiber (2003), une définition aristotélicienne de celle-ci ne permet pas de rendre compte de la diversité des configurations rencontrées en discours.

⁴ En revanche, elle peut rappeler certaines pédagogies de l'enseignement des langues pour lesquelles la maîtrise du code passe par la répétition de phrases visant à l'assimilation de structures clés.

Selon le linguiste, c'est par le biais d'une redéfinition prototypique que le concept semble pouvoir conserver quelque consistance. En un sens, la phrase canonique semble bien avoir anticipé cette solution, plus ou moins consciemment. Ce faisant, elle compense par ailleurs un problème épistémologique pointé par Culioli (1990) et Milner (1989), à savoir le fait que la langue est, pour le linguiste, à la fois un objet d'investigation et un outil descriptif.

Références

Références bibliographiques

- CULIOLI, Antoine (1990). *Pour une linguistique de l'énonciation t.I – Opérations et représentations*. Paris : Ophrys.
- DENIS, Delphine & SANCIER-CHÂTEAU, Anne (1994). *Grammaire du français*, Paris : Livre de Poche.
- GAUTIER, Antoine (2006). *Unité et discontinuité : une approche épistémologique et systématique de la phrase*. Thèse de doctorat, Paris IV.
- GERMAIN, Claude & SÉGUIN, Hubert (1998). *Le point sur la grammaire*. Paris : CEC, CLE international.
- GREENBERG, Joseph H. (Ed) (1966). *Universals of Language*, Second edition, Cambridge : The MIT Press.
- KLEIBER, Georges (1990). *Sémantique du prototype*, Paris : P.U.F.
- KLEIBER, Georges (2003). Faut-il dire « adieu » à la phrase ?, *L'Information grammaticale*, 98.
- LAMBRECHT, Knud (à paraître, 2007). Contraintes cognitives sur la syntaxe de la phrase en français parlé, in *Actes du Colloque Modèles syntaxiques, Bruxelles, 6-8 mars 2003*, Van Raemdonck, Dan (Ed). s.l.n.d.
- LE GOFFIC, Pierre (1994). *Grammaire de la phrase française*. Paris : Hachette.
- MILNER, Jean-Claude (1989). *Introduction à une science du langage*. Paris : Seuil.
- RIEGEL, Martin, PELLAT, Jean-Christophe & RIOUL, René (1994). *Grammaire méthodique du français*. Paris : P.U.F.
- ROSCH, Eleanor (1973). Natural categories, *Cognitive psychology*, 4, San Diego : Academic press.
- WILLEMS, Dominique (2001). Linéarité et métalangue : de la phrase type aux types de phrases, in *Métalangage et terminologie linguistique, Actes du colloque international de Grenoble (Université Stendhal – Grenoble III, 14 16 mai 1998)*, Colombat, Bernard & Savelli, Marie (Eds). Leuven-Paris-Sterling : Peeters.

Références aux sites Internet

- ULLAND, Harald, Les titres rédactionnels phrastiques et leur analyse automatique en prédicat et arguments, <http://www.duo.uio.no/roman/art/rf-16-02-2/fra/ulland.pdf>, (2002).

SUR LA VALEUR "PROTOTYPANTE" DU PRÉSENT DE L'INDICATIF

Rose-Marie GERBE

rose-marie.gerbe@u-grenoble3.fr

Equipe TRAVERSES 19/21

38000 Grenoble

Abstract : In some written or oral utterances, the present tense does not refer to any specific situation, whether simultaneous to the speech act or not. It does allow a "prototypical" event to be referred to, through a co-text and a context involved in this "dis-actualization". This event is unreal and exemplary and can be realized in many different ways; that is the signification of "prototyping modality" of the present tense.

La liste des emplois contextuels du présent de l'indicatif (PR) est longue : historique, déictique, omnitemporel, atemporel, futur, hypothétique, etc. Ces emplois ont été expliqués traditionnellement selon trois angles d'étude : la théorie temporelle, dans laquelle on considère que les emplois contextuels du PR ne seraient que « déviants » par rapport à sa valeur première déictique ; la théorie aspectuelle, dans laquelle le PR est envisagé essentiellement comme une forme sécante ; la théorie de « zéro morphème » ou de tiroir neutre : le PR ne permettrait pas d'attribuer une couleur temporelle au procès dénoté et serait alors la catégorie non marquée du système verbal. Le signifié commun de tous les emplois du PR ne saurait être temporel, puisque l'effectivité de la référence n'est pas toujours pertinente (voir Gerbe, 2006) ; serait-il alors « modal » ? Existe-t-il des contextes singuliers où le PR relève d'une certaine modalisation ? Dans certaines occurrences en effet, le PR semble résister à la référence temporelle et à l'actualisation (passage d'un concept en une représentation particulière de sujets parlants). Les analyses traditionnelles du PR ne seront par exemple d'aucun secours pour les énoncés français, à caractère modal, qui nous occupent : trois types d'énoncés écrits (exemples philosophiques, scénarios narratifs et textes de loi) et deux types d'énoncés oraux (discours procédural et histoire drôle), tous caractérisés par une formulation systématique au PR et par l'effacement du producteur de l'énoncé¹. Il faudra envisager une autre problématique, qui permettrait d'attribuer au PR non seulement une valeur d'assertion (d'un point de vue du mode), mais aussi une valeur de mise en attente (d'un point de vue de la modalité). On résumera la problématique ainsi : le PR peut dénoter un événement qui n'a pas d'ancrage et qui représente un prototype d'événement².

Si l'étendue des corpus peut paraître insolite, elle est néanmoins maîtrisée : l'analyse modale du PR dépasse strictement le cadre du texte et des indices convergents (tant du côté du contexte que du cotexte) permettent de rapprocher des énoncés si divers. Ce qui nous importe est la façon dont le PR interagit avec les autres éléments pour former cette modalisation prototypante. Les phénomènes énonciatifs seront à étudier, puisqu'ils sont liés au fait que lecteur et scripteur ne sont pas clairement désignés dans nos exemples ; le contexte aura une place prépondérante lors de la réception, notamment pour les valeurs illocutoires (injonctive ou prévisionnelle) qui viennent s'ajouter à la prototypie (référence non actualisée et visée non unique du référent).

¹ Nous avons relevé parmi de nombreuses occurrences les énoncés symptomatiques des phénomènes remarqués (voir les exemples 1 à 8), pour ne traiter ici que de l'essentiel.

² Précisons que le prototype d'événement (en tant que procès décrit par le verbe conjugué) est l'exemplaire d'une classe ; il admet un ensemble de réalisations en saisie paradigmatique.

1. Modalisation et référence

Les énoncés que nous allons présenter et analyser ne sont pas les seuls à avoir une teneur modale. Outre les énoncés oraux avec du PR modal (comme dans « On boit un coup ? »), il y a quelques énoncés écrits qui dépassent le simple aspect itératif³. Ainsi, Mellet relève, dans un récit fictionnel avec contexte passé, un PR qui offre l'image d'une journée prototypique (2001 : 30) ; dans le passage cité, l'évènement décrit peut être la quintessence de l'ensemble des évènements possibles. Parce que le PR opère un lissage de la mise en perspective (il ne permet pas de différencier le premier plan de l'arrière-plan) et un aplatissement de la chronologie, il est utilisé pour transformer les évènements en des « images prototypiques » (2001 : 34). Avec le PR de narration donc, l'effet de prototypie peut apparaître. C'est ce que Rebotier a également noté, en repérant un autre exemple de PR de narration prototypant : il s'agit d'une séquence fictionnelle, dans laquelle est décrite l'entrée type d'un professeur dans une classe, rédigée au *Präsens* allemand (2005 : 218). Le tiroir fait alors vivre la scène décrite comme si elle était un prototype de toutes les scènes semblables (2005 : 426). Le *Präsens* allemand peut donc – comme le PR français – décrire un évènement représentant un pur schéma abstrait, un patron d'évènement. Il nous faut préciser que le PR de narration a rarement cette valeur prototypante ; dans nos exemples, le PR, de par son interaction avec les présentatifs et les marques génériques notamment, est toujours prototypant.

1.1. La notion de « prototype d'évènement » : il y a prototype et prototype

La problématique décrite ici est différente de la sémantique du prototype, dans laquelle le prototype est le meilleur exemplaire d'une catégorie en tant qu'exemplaire le plus fréquemment donné comme tel (Kleiber, 1990 : 72). Illustrons cette différence par un syntagme verbal donné par Kleiber lui-même : « le scénario *aller au restaurant* rassemble des instances constituées d'une suite d'actions jugées prototypiques et d'autres qui comprennent des éléments inhabituels, non attendus. » (1990 : 103-104). Dans la sémantique du prototype, « prendre un taxi, se rendre dans un restaurant gastronomique et commander un menu » par exemple est le prototype d'*aller au restaurant*, alors que « partir à Londres pour manger dans un *Fish and Chips* » n'est pas prototypique ; dans notre approche, *aller au restaurant* est un prototype d'évènement, qui peut se réaliser par une multitude de procès (dont ceux précédemment cités). Dans un scénario romanesque (voir plus bas), ce phénomène est particulièrement évident : si l'auteur écrit « Julien va au restaurant. », le texte définitif pourra ressembler à « Julien quitta son modeste appartement de la rue des Martyrs et se dirigea vers Le Café de Paris, où il avait ses petites habitudes. ». C'est principalement le PR qui provoque cette prototypie : il ne sert pas à décrire un état de fait, anecdotique ou imaginaire, mais à décrire un évènement susceptible de se produire ou de s'être produit, dans des circonstances spatio-temporelles à spécifier.

1.2. Localisation spatio-temporelle ?

Nous parlions plus haut de la non-pertinence de l'angle temporel pour analyser certains PR. En effet, dans nos énoncés, les indications spatiales ou temporelles sont soit absentes soit très vagues : les phénomènes d'actualisation et d'ancrage sont effacés. Par opposition au PR de narration par exemple, qui nécessite un contexte passé évident et une référence spécifique (à un fait du monde réel ou virtuel), le PR ne peut s'appuyer sur le cotexte pour référer spatio-temporellement. C'est du côté de la théorie praxématique que nous devons regarder : le PR a la capacité de saisir l'évènement au niveau émergent de l'actualisation (dernier niveau selon

³ Ce sera la seule mention de l'aspect dans cet article ; il est clair que l'aspect inaccompli du PR a un lien avec notre valeur prototypante, mais nous ne pouvons être exhaustive ici.

Barbérès, 2004 : 207-208). Il peut par conséquent dépasser l'actualisation pure et référer à un événement qui n'est pas complètement actualisé, et qui par ailleurs n'est pas ancré dans le temps (ni dans la situation d'énonciation, ni dans une autre époque temporelle). Dans un certain contexte, le recours au PR provoque l'effacement de la temporalité, en référant à une situation imaginaire « pseudo-déictique », et laisse libre l'actualisation de l'événement dans le monde réel. L'auteur décrit une scène qui a des qualités intrinsèques (qui permettent donc au lecteur de la visualiser) et qui représente pourtant l'exemplarité, que l'on repère principalement par le manque d'ancrage spatio-temporel. Cette lecture ne peut être évidente qu'après l'étude détaillée des corpus ; c'est l'ensemble du passage et du contexte de production et de réception qui doit être envisagé.

2. Le PR et les autres marques de prototypie

Trois éléments sont fondamentaux dans la modalisation prototypante : le contexte d'apparition, le cotexte et la formulation au PR. La notion d'« appareil formel »⁴ pourra nous être utile pour montrer que le PR est l'élément central dans ce système.

2.1. Le contexte d'apparition

Les cinq genres d'énoncés que nous rapprochons ont au moins deux caractéristiques énonciatives identiques : le contexte extralinguistique permet d'une part l'effacement de la relation énonciateur / destinataire et d'autre part la mise en place d'une valeur illocutoire.

- Les exemples philosophiques : valeur prototypante du PR

Ce premier corpus pourra nous servir de « guide », puisque que le PR y a une valeur prototypante absolue, sans autre valeur pragmatique :

1. Je lance une pierre. Elle traverse mon jardin. Elle devient pour un moment un bolide confus et redevient pierre en tombant sur le sol à quelque distance. (Merleau-Ponty, Maurice, *Phénoménologie de la perception*, 1945, Paris : Gallimard, p. 310)
2. J'ai rendez-vous avec Pierre à quatre heures. J'arrive en retard d'un quart d'heure [...] Je regarde la salle, les consommateurs et je dis : « Il n'est pas là. » (Sartre, Jean-Paul, *L'Être et le Néant*, 1943, Paris : Gallimard, p. 43-44)

Dans les exemples philosophiques, une convention sera établie : formellement, les marques non spécifiques (du type *une pierre* et *Pierre*) laissent en attente la validation du procès et les indices de fictionnalité (du type *voici* et *supposons*) vont introduire des séquences en décalage avec les séquences théoriques discursives ; pragmatiquement, le lecteur perçoit que l'intérêt n'est pas de référer à une anecdote personnelle – avec un *je* embrayeur – mais de référer à un événement susceptible de se produire.

- Les scénarios narratifs : valeur prototypante et prévisionnelle du PR

Ce second corpus contient des procès au PR référant à des événements prototypiques, eux-mêmes en instance de réécriture :

3. Mathô les observe, Narr'Havas les fait s'arrêter, s'avance seul, fait signe qu'il veut parler à une sentinelle, remet sa lance et son cheval et s'avance seul à pied. (Flaubert, Gustave, scénario de *Salammô*, repris dans Gagnebin, Bernard, *Flaubert et « Salammô » : genèse d'un texte*, 1992, Paris : Presses Universitaires de France, p. 67)⁵

⁴ L'appareil formel a ici un sens dérivé de la théorie de Benveniste : un ensemble de traits linguistiques congruents, prévisibles ou obligatoires, c'est-à-dire les marques formelles surreprésentées (cf. Philippe, 2002 et 2005).

⁵ Ce passage sera transformé dans le texte définitif en « Tout à coup, Narr'Havas marcha résolument vers une sentinelle. Il jeta son épée, sa lance, ses javelots, et disparut au milieu des Carthaginois. » (*Salammô*, 2001, Paris : Flammarion, p. 274).

4. C'est alors qu'en sortant des bâtiments par les escaliers noirs, il rencontre les Maheu et les Levaque qui arrivent. (Zola, Emile, scénario de *Germinal*, repris dans Becker, Colette, *Zola : le saut dans les étoiles*, 2002, Paris : Presses de la Sorbonne Nouvelle, p. 323)⁶

Les scénarios sont des écrits éminemment privés, à rapprocher des écrits d'archive : « leur réception se fait sur le mode de la "consultation" et non de la "communication" » (Philippe, 2002 : 26). Le PR participe à la construction de l'histoire, à la construction du roman, qui va devenir, après plusieurs étapes de réécritures, un récit dans lequel le destinataire peut être déterminé. Plusieurs opérations marqueront la transformation de l'« avant-texte » en texte :

- le passage de groupes nominaux sans déterminant, qui relèvent d'un renvoi à une notion, à des groupes nominaux instanciés⁷, avec une détermination précise ;
- le passage de nombreux substantifs déverbaux, de participes présents et de verbes au PR à des verbes plus précis et des tiroirs permettant de différencier deux niveaux d'énonciation, l'un marqué par l'imparfait (arrière-plan), l'autre par le passé simple (premier plan) ;
- le passage d'une quasi absence de circonstant spatio-temporel à une émergence de connecteurs, de circonstants de lieu, de circonstants de temps...

Etant donné que l'écrivain sait que personne ne lira sa production, que ce ne sont que des « brouillons » au plein sens du terme, il peut se permettre d'employer une syntaxe purement notationnelle, en omettant par exemple les déterminants et les sujets des verbes. Ce style épuré coordonné à l'usage du PR sera un indice supplémentaire pour l'analyse prototypante des scénarios : l'écrivain « réfère » à une histoire embryonnaire.

- Les textes de loi : valeur prototypante et injonctive du PR

Ce troisième corpus est lié aux notions de généricité et d'exemplarité, qui s'accommodent avec la prototypie événementielle :

5. Les époux s'obligent mutuellement à une communauté de vie. / La résidence de la famille est un lieu qu'ils choisissent d'un commun accord. (article 215 du *Code civil 2006*, Paris : Dalloz)
6. Le divorce peut être demandé conjointement par les époux lorsqu'ils s'entendent sur la rupture du mariage et ses effets en soumettant à l'approbation du juge une convention réglant les conséquences du divorce. (article 230 du *Code civil 2006*, Paris : Dalloz)⁸

Dans les textes de loi, la visée non communicationnelle peut aussi être convoquée : il y a une « distance » importante entre les deux acteurs du message. L'effacement énonciatif (on *fait comme si* l'énoncé n'avait pas été écrit par quelqu'un et pour quelqu'un) n'est possible que parce qu'il existe un « pacte de lecture » : le lecteur, qu'il soit juriste ou non, sait que la loi est prescriptive ; l'auteur n'est pas obligé de marquer explicitement son discours pour le doter de visée illocutoire. C'est ce qui expliquerait l'absence du mode verbal impératif et l'absence de mention d'un destinataire direct dans ce genre d'écrit.

⁶ Ce passage donnera lieu à « le jeune homme [...] s'approchait, lorsqu'il rencontra une nouvelle bande de charbonniers, qui arrivait à la fosse. » (*Germinal*, 2000, Paris : Flammarion, p. 74).

⁷ L'instanciation désigne, dans la théorie énonciative, l'opération qui consiste à saturer les positions d'une structure prédicative au moyen de notions, c'est-à-dire de signifiés lexicaux. Un groupe nominal instancié sera ainsi un groupe nominal ancré dans un groupe verbal, dans une prédication.

⁸ Cet article remplace l'article 281 du *Code civil* de 1804 (repris dans Acolas, Emile, *Manuel de droit civil : commentaire philosophique et critique du Code Napoléon*, 1874, Paris : Germer-Baillière, tome I, p. 265), rédigé au futur : « Les époux se représenteront ensemble, et en personne, devant le président du tribunal civil de leur arrondissement, ou devant le juge qui en fera les fonctions, et lui feront la déclaration de leur volonté, en présence de deux notaires amenés par eux ».

Les deux valeurs attribuées au PR (prototypant et injonctif) relèvent également du contexte, et en particulier des pronoms indéfinis et des structures verbales. La plupart des articles de loi est effectivement rédigée avec un sujet « indéterminé » singulier et notamment un pronom indéfini. De plus, les rédacteurs varient le type de structure verbale, selon que l'injonction est plus ou moins présente. Trois grands types de constructions se manifestent : les auxiliaires modaux ou assimilés (*pouvoir, devoir, être tenu de, avoir la faculté / possibilité de...*), les formes impersonnelles (*il faut que, il est permis de, il convient de...*) et les formes passives (*il est demandé, telle chose se fait...*). Avec la forme passive, très courante, « l'essentiel se reporte sur l'objet » et « les opérateurs importent moins que l'opération » (Cornu, 2000 : 328). Il convient de faire parler la loi plutôt que le législateur, la troisième personne plutôt que la première, voire de taire le nom de celui qui parle, laissant ainsi la règle s'imposer d'elle-même, comme émanant de la volonté commune. On rejoint par conséquent la primauté de l'effacement énonciatif, qui est un corollaire de l'atténuation de l'injonction par la rédaction au PR, qui remplace depuis le XX^e siècle la rédaction au futur (voir note 8).

- Les descriptions procédurales : valeur prototypante du PR ?

Ce quatrième corpus contient des PR qui nous semblent également prototypants :

7. Vous devez subir certains concours alors vous passez votre concours si vous êtes apte vous rentrez à l'école de gendarmerie [...] à la sortie de ces huit mois vous avez donc un classement et vous choisissez une place... (discours procédural issu des *Corpus de français parlé contemporain établis par les étudiants de maîtrise*, 1996, sous la direction de Marie Savelli, Université Stendhal-Grenoble III)

Dans les descriptions procédurales à l'oral, le PR apparaît en co-occurrence avec des formules introductives (du type *je t'explique* ou *ce qu'il faut faire*) et des pronoms de 2^e personne « embrayeurs virtuels » (voir 2.2.). L'évènement n'est pas raconté comme fictionnel et historique (ce qui serait le cas avec le PR de narration, très commun à l'oral) mais devient prototypique : le schéma d'action est similaire dans tous les cas. Dans ces énoncés procéduraux, « [l]'énonciateur expose simplement une probabilité, décrochée totalement du moment de l'énonciation. La validation du procès n'a aucune importance. » (Delveroudi, 2002 : 20-21). Autrement dit, l'évènement décrit reste en attente de validation, comme dans les énoncés écrits précédemment cités.

- Les histoires drôles : valeur prototypante du PR ?

Ce dernier corpus est d'emblée placé dans un monde virtuel et « désactualisé » :

8. Un type conduit sur une route de montagne. Dans un virage, il croise une femme dans une superbe Mercedes. La fille ouvre sa vitre et crie « Attention gros porc ! ». Le type ouvre sa vitre à son tour et crie « Petite conne ! ». Et il continue sa route. Un peu plus loin, il percute un cochon qui broute sur le bas-côté, sort de la route, tombe dans le ravin et meurt. Moralité : les hommes feraient bien, de temps en temps, d'écouter ce que les femmes disent. (énoncé relevé par nos soins).

Les histoires drôles, bien que récits fictifs, sont inimaginables au passé simple ou à l'imparfait : il y a une corrélation exclusive entre le PR et l'histoire drôle. Le propre de la blague étant de raconter un évènement mais de ne prétendre en rien que cet évènement se soit passé, le PR est de mise puisqu'il n'inscrit pas la représentation des évènements en réalité (Bres, 2005 : 31). On se rapproche bien d'une valeur prototypante du PR.

En conséquence, dans nos cinq types d'énoncés, le procès est posé comme non validé mais valable par n'importe quel sujet, à n'importe quel moment :

« [d'un côté] valable n'implique pas que cette relation sera nécessairement validée ou, au contraire, non validée (on imagine toutes les éventualités) ; d'un autre côté, valable signifie "qui peut être validée, qui est donc validée

pourvu qu'il y ait un énonciateur pour effectuer l'opération". » (Culioli, 1980 : 185).

La notion de « désactualisation » est liée à celle d'atemporalité et de validation : avec le présent, étant donné qu'il invalide les repères, la relation sera posée comme validable à la fois pour T_0 , pour T_{-1} , pour T_{+1} , etc. Le procès peut être présenté d'une manière réaliste ; ce qui importe est son exemplarité même et la mise en attente de son actualisation. Cette mise en attente revêt, en particulier à travers un contexte précis, d'une teneur prévisionnelle – dans le cas où l'énoncé est un outil pour construire un énoncé plus standard – ou d'une teneur injonctive – dans le cas où l'énonciateur est par définition autoritaire.

2.2. Le cotexte : valeurs des pronoms personnels et des « prénoms »

Nous nous limiterons à deux indices qui interagissent avec l'usage du PR, notre objectif étant de relever des formes faisant faisceau et donnant lieu à une interprétation similaire (en l'occurrence une lecture non temporelle et prototypante) : les pronoms personnels embrayeurs et les noms propres de personne.

- Le pronom de 1^{re} personne

Dans les exemples philosophiques, la valeur du pronom de 1^{re} personne est particulièrement intéressante. Le pronom *je* ne renvoie pas spécifiquement à l'auteur mais renvoie à une entité virtuelle (voir exemples 1 et 2). Les personnes du discours utilisées ne proposent pas de prise en charge énonciative franche : ce sont des *je* génériques qui appellent sans cesse le lecteur à une « identification participante » (Temmar, 2003 : 167) ; chaque lecteur va s'identifier à l'acteur de l'évènement dénoté. On utilisera l'expression d'« embrayeur virtuel » : *embrayeur* puisque le pronom personnel garde manifestement un certain effet de « présence », et *virtuel* puisque le *je* renvoie à tout être humain susceptible de ramener son expérience au prototype de l'évènement. La référence du pronom est ainsi « non spécifiée », tout comme la référence des procès décrits au PR.

Ces exemples n'ont de pleine pertinence philosophique que parce que la référence des embrayeurs reste purement virtuelle : *je* ouvre une classe de locuteurs potentiels. Le choix de l'embrayeur virtuel *je* est donc lié à une démarche argumentative et à une volonté d'abstraction et de généralisation. Les philosophes choisissent de raconter un évènement dans sa singularité, tout en lui donnant un minimum de précision et un maximum de « pseudo-référence ». Nous assistons au même phénomène dans les descriptions procédurales : les pronoms de 2^e personne peuvent y être aussi qualifiés d'embrayeurs virtuels (le *tu* et le *vous* ne renvoient pas uniquement à l'interlocuteur).

- Les « prénoms »

Théoriquement, le prénom est un nom propre et réfère à un être particulier, inconfondable avec d'autres, qui existe à travers une identité spécifique. Ce principe de référence semble absent des exemples de grammaire, de définition, de philosophie : *Pierre*, *Paul* et *Marie* y sont couramment utilisés et sont plus proches alors des pronoms indéfinis que des noms propres, d'où la possible locution « pronoms propres ». Le « pronom propre » est un usage du prénom qui permet de laisser en suspens la référence, ce qui est contraire en principe au qualificatif « nom propre ». La locution « pronom propre » permet de rendre compte que les prénoms tiennent lieu à la fois de noms propres (chaque renvoi est évident pour le lecteur) et de pronoms indéfinis (la personne à laquelle ils renvoient n'est pas spécifique). Par exemple, dans les œuvres philosophiques, *Pierre* (prénom utilisé le plus couramment) a l'avantage de personnifier le sujet de l'évènement, tout en disant au lecteur « Pierre n'existe pas ». Les prénoms sont choisis délibérément pour leur neutralité, leur universalité ; ils bloquent toute inférence sur le milieu social, l'âge, etc. du référent. Dans les histoires drôles également, les prénoms sont utilisés pour référer à une personne imaginaire et non spécifique ; à l'inverse,

ils sont souvent connotés : on utilisera *Marie-Chantal* pour une bourgeoise, *Ginette* pour une ménagère, *Kevin* pour un adolescent⁹. Les « pronoms propres » sont donc un indice fort de prototypie ; le lecteur perçoit que la saynète n'est pas actualisée, qu'il ne doit pas chercher la référence exacte de *Pierre* ou de *Ginette*.

2.3. Le PR par rapport aux autres tiroirs

Le PR est volontairement choisi dans les énoncés reproduits ici pour ses valeurs neutralisantes : il permet la « désactualisation », et, partant, l'exemplarité, de l'évènement.

- Le PR à la place du FS (exemples 5 et 6)

Dans les textes de loi, le futur étant trop marqué temporellement et illocutoirement, le PR a été préféré, pour un droit moins prohibitif et plus proche d'une description que d'une prescription. Ainsi, pour le Code pénal, qui a longtemps gardé les futurs simples, la structure des articles est maintenant quasiment identique : « *le fait de faire ceci est puni de cela* » (au lieu de « Celui qui fera ceci sera puni de cela »). En partant du principe que si les dispositions sont écrites comme une narration d'évènements fictifs, elles éviteront le problème de la gestion du temps, les rédacteurs rédigent les Codes au PR. Le PR est alors destiné à suggérer la vérité permanente des principes invoqués, en mettant ceux-ci à l'abri de tout contexte historique. En plus de sa simplicité morphologique, qui est un atout supplémentaire pour la lisibilité des textes de loi, le PR a également l'avantage de renvoyer à des actes coupés de toute temporalité, de toute chronologie ; il est dans un mouvement perpétuel qui « colle » à une société mouvante, tout en respectant le caractère intrinsèquement abstrait des dispositions juridiques.

- Le PR, seul tiroir « prototypant » ?

Le PR est probablement le seul tiroir qui accepte une valeur prototypante. Les autres tiroirs présentent le procès comme actualisé dans le passé (qu'il soit réel ou virtuel) ou dans le futur (hypothétique) : l'évènement est alors envisagé comme unique et spécifique. Pour qu'un énoncé acquière cette valeur prototypante, il faut que la rédaction soit entièrement au PR. A l'oral par exemple, une histoire qui commence par des verbes au PR sera ordinairement analysée comme une histoire drôle ; si le blagueur utilise les tiroirs du passé, les destinataires pourront interpréter le message comme une histoire vraie. Il suffit d'introduire dans ce type de narration un seul passé composé pour que la réception de l'interlocuteur soit perturbée et qu'il se demande si « c'est du lard ou du cochon » (Bres, 1998 : 132). Avec le PR, le narrateur n'attribue pas clairement à l'évènement risible une singularité.

3. Conclusion

Les phénomènes d'effacement énonciatif, de non-ancrage et de modalisation sont liés à l'appareil formel, dont le PR est ici pivot. Le PR dans les énoncés « désactualisés » est « textualisant » (c'est-à-dire qu'il ne reçoit cette valeur que dans le cas d'une série de procès ; un énoncé isolé n'a aucune chance d'être interprété comme opérant un changement modal) et « non actualisant » (il suspend l'effectivité de la prédication) (Philippe, 2005 : 83). Le PR permet de maintenir la séquence dans une narration d'évènements, en rendant ces mêmes évènements extérieurs à toute référence temporelle et à toute réalité. Fonctionnant avec des embrayeurs virtuels, des « pré-noms » à valeur de pronoms indéfinis, le PR opère un décrochage modal et le lecteur perçoit l'évènement comme non ancré dans l'histoire. Le PR prototypant renvoie à un évènement susceptible de se (re)produire dans un univers réel :

« Le présent, par lui-même, n'inscrit pas la représentation des évènements en réalité temporelle, pas plus dans le passé que dans une autre époque, et donc

⁹ L'équivalent de *Pierre* des exemples philosophiques pourrait être *Toto* dans les histoires drôles, mais son panel de réalisations nous paraît plus étroit (Toto s'apparentant souvent à un enfant naïf).

tout simplement ne l'inscrit pas en réalité : la représentation n'est pas à rapporter à un référent. » (Bres, 2005 : 31).

Le PR explicite le fait que la référence du procès dénoté ne soit pas la principale volonté de l'énonciateur. Le repérage de l'évènement peut être à chaque lecture différente ; c'est ce qu'on entend par « non validé mais validable » (voir plus haut). Dans les exemples cités, le PR donne lieu à un décrochage modal – dû au statut ambigu de certaines marques (par exemple des pronoms personnels) – et à la transformation d'un procès fictif en un procès prototypique. C'est cette transformation qui est souhaitée par l'auteur, qui veut faire participer le lecteur par une appropriation de l'univers créé. Non seulement il va s'approprier le *je* utilisé, mais aussi le PR, qui n'a pas la stricte valeur déictique qu'on veut lui attribuer. Notre travail se situe donc à la jonction entre une étude formelle et une étude discursive, voire pragmatique. La valeur prototypante est l'effet d'interactions avec d'autres éléments cotextuels : elle est le résultat d'un jeu subtil entre plusieurs facteurs contextuels, qui sont à rassembler sous la notion d'appareil formel, le facteur principal étant le recours au PR.

Références

- BARBÉRIS, Jeanne-Marie (2004). Coénonciation et actualisation : la textualisation en *idem*, in *L'analyse du discours dans les études littéraires*, Amossy, Ruth & Maingueneau, Dominique (Eds). Toulouse : Presses Universitaires du Mirail, 201-212.
- BRES, Jacques (1998). De l'alternance temporelle passé composé / présent en récit oral conversationnel, in *Variations sur la référence verbale*, Borillo, Andrée, Vetters, Carl & Vuillaume, Marcel (Eds). Amsterdam/Atlanta : Rodopi B.V. (Cahier Chronos 3), 125-136.
- BRES, Jacques (2005). Le présent de l'indicatif en français : de quelques problèmes, et peut-être de quelques solutions, in *Du présent de l'indicatif*, Despierres, Claire & Krazem, Mustapha (Eds). Dijon : Université de Bourgogne, 27-51.
- CORNU, Gérard (2000). *Linguistique juridique*. Paris : Montchrestien-E.J.A.
- CULIOLI, Antoine (1980). Valeurs aspectuelles et opérations énonciatives : l'aoristique, in *La notion d'aspect*, David, Jacques & Martin, Robert (Eds). Metz/Paris : Centre d'analyse syntaxique-Klincksieck, 181-193.
- DELVEROUDI, Rhéa (2002). A propos d'une des valeurs du présent de l'indicatif français : la valeur injonctive, *Revue de sémantique et pragmatique*, 11 : 7-25.
- GERBE, Rose-Marie (2006). *Le présent prototypant : contribution à l'étude des appareils formels du français écrit*, thèse de doctorat sous la direction de Gilles Philippe. Grenoble : Université Stendhal-Grenoble III.
- KLEIBER, Georges (1990). *La sémantique du prototype : catégories et sens lexical*. Paris : Presses Universitaires de France.
- MELLET, Sylvie (2001). Valeur aspectuelle du présent : un problème de frontière, in *Le présent en français*, Le Goffic, Pierre (Ed). Amsterdam/Atlanta : Rodopi B.V. (Cahier Chronos 7), 27-39.
- PHILIPPE, Gilles (2002). L'appareil formel de l'effacement énonciatif et la pragmatique des textes sans locuteur, in *Pragmatique et analyse des textes*, Amossy, Ruth (Ed). Tel Aviv : French Department, 17-34.
- PHILIPPE, Gilles (2005). Existe-t-il un appareil formel de la fiction ?, *Le français moderne*, LXXXIII, 1 : 75-88.
- REBOTIER, Aude (2005). *Les Temps verbaux en allemand et en français et les phénomènes aspectuels : comparaison de deux systèmes*, thèse de doctorat sous la direction de Jean-François Marillier. Grenoble : Université Stendhal-Grenoble III.
- TEMMAR, Malika (2003). Effets de présence et ostension : une approche discursive, *Studia Phaenomenologica*, III, 3/4 : 163-184.

ANALYSE TEXTUELLE, ANALYSE DE GENRES : QUELLES RELATIONS, QUELS INSTRUMENTS ?

Mathilde GONÇALVES, Florencia MIRANDA
matilde.goncalves@fcsh.unl.pt, fmiranda@arnet.com.ar
 CLUNL, Av. De Berna, 26C
 1069- 061 Lisboa
 Portugal

Abstract : The main goal of this work is to present and discuss the issues raised by the distinction and instrumentalization of two different approaches: “textual analysis” and “genre analysis”. Analysis instruments and specific examples from a project being developed by the Linguistic Centre of Universidade Nova de Lisboa will be presented.

1. Introduction

Le présent travail s’intègre dans un Projet de Recherche qui est développé par le Centre de Linguistique de l’Université Nouvelle de Lisbonne. La problématique générale de ce projet s’intéresse aux relations entre « genres textuels et organisation de la connaissance » et se situe dans la perspective de la Théorie du Texte. Comme domaine de spécialisation dans les sciences du langage, nous envisageons la théorie du texte, dans laquelle nous nous situons, comme champ d’étude vers lequel convergent différentes perspectives sur les textes et les discours, et dans laquelle la notion de genre occupe une place centrale (cf. Adam, 1999 ; Bronckart, 1997 ; Maingueneau, 1998 ; Rastier, 2001).

Étant donné que l’objectif de ce Projet est d’observer les imbrications possibles entre les différentes modalités d’organisation de la connaissance et les divers genres de textes¹, un instrument d’analyse pertinent pour la caractérisation des genres en question est essentiel. Ainsi, dans la première étape de la recherche, étape qui est en phase de conclusion, nous nous sommes centrés sur la problématique (théorico-méthodologique) de la distinction entre deux approches – d’une part, l’analyse textuelle et d’autre part, l’analyse de genres textuels. Dans notre travail, nous aborderons ce premier aspect de la recherche dans le but de montrer comment nous concevons ces deux approches.

Dans un premier temps, nous présenterons succinctement notre instrument d’analyse et les catégories analytiques qui lui sont associées. Puis, nous exemplifierons la façon dont cet instrument est utilisé à travers l’analyse d’un des genres étudiés pour le projet : l’annonce publicitaire. Pour ce, nous nous proposons d’observer un des multiples aspects faisant partie de l’organisation interne des textes et des genres : l’articulation entre unités textuelles de nature sémiotique différente.

2. Analyse textuelle et analyse des genres: un problème théorico-méthodologique

Dans le cadre de notre recherche sur la relation entre genres et organisation de la connaissance, nous nous sommes fixé comme but la caractérisation de quelques genres textuels associés à diverses activités langagières : prises de notes, affiches politiques,

¹ Le choix du terme « genres de textes » au lieu de « genres du discours » fait référence aux discussions sur le sujet faites par des auteurs comme Rastier et Bronckart. Cf. à titre d’exemple Bronckart (2004 : 102).

annonces publicitaires et écriture fragmentaire. La première difficulté à laquelle nous avons été confrontés est d'ordre théorico-méthodologique : comment décrire les genres ? Avec quels instruments ? De fait, les modèles de description proposés par les théoriciens situés dans diverses perspectives convergentes, comme Adam, Bronckart et Rastier, sont explicitement présentés comme étant des instruments d'analyse de « texte » et non pas de « genre ». La question qui se pose est de savoir si ces instruments pourraient être utilisés pour la description des genres.

Pour répondre à cette question, nous assumons comme première hypothèse de travail que « analyse de textes » et « analyse de genres » sont deux démarches d'ordre différent, qui s'imbriquent mutuellement.

L'analyse textuelle est une activité qui a une incidence sur les textes empiriques, attestés, c'est-à-dire, sur des objets sémiotiques singuliers. Ainsi lorsqu'on réalise une analyse textuelle, l'objectif est, au départ, de rendre compte des spécificités de ces textes dans ce qu'ils ont d'individuel. En contrepartie, l'analyse de genres est une activité qui vise à identifier les traits spécifiques d'un genre déterminé. Dans ce sens, il s'agit d'une opération qui prétend montrer les caractéristiques qui permettent la reconnaissance de l'inscription générique des textes. Pour identifier ces traits ou caractéristiques, il faut effectuer une analyse textuelle sur un corpus de textes appartenant au même genre. Remarquons que dans ce cas l'analyse de textes n'est pas une fin en soi mais bien une étape de distinction des caractéristiques singulières et des caractéristiques récurrentes.

En conséquence de quoi, nous avons voulu vérifier si les instruments développés pour une analyse textuelle peuvent se montrer pertinents pour une analyse de genres. Ceci nous a amenés à formuler l'hypothèse d'une possible identification de deux catégories analytiques, qui nous permettent de rendre compte des phénomènes associés à chaque approche : « paramètres génériques » et « mécanismes de réalisation textuelle ». Les « paramètres » fonctionnent, dans la production textuelle, comme normes pour la construction d'un texte inscrit dans le genre en question. Ces « paramètres » sont des données que le chercheur obtient de l'analyse de genre. Les « mécanismes » sont des actualisations des paramètres et ont trait à la gestion des ressources semio-linguistiques concrétisée dans chaque texte empirique. Ces « mécanismes » sont donc obtenus à partir de l'« analyse de textes ».

Dans cette étape de notre travail, nous nous sommes centrés sur la construction d'un instrument d'analyse. Nous nous sommes donc basés sur le modèle de l'architecture textuelle telle que l'interactionisme socio-discursif (ISD) la conçoit (cf. Bronckart, 1997) – tout texte est constitué de trois niveaux organisationnels (infrastructure générale, mécanismes de textualisation et mécanismes de prise en charge énonciative), qui se superposent les uns aux autres et interagissent entre eux.

Le positionnement épistémologique général de ce modèle et les grandes lignes qui le composent n'ont pas été remis en cause. Cependant, après un travail d'observation critique et d'application empirique, nous avons proposé l'introduction et, éventuellement, la reformulation de certains éléments. Pour ce, en plus de notre travail sur les textes, les contributions d'autres théoriciens ont été essentielles, notamment, Adam (1999 et 2005), Bernárdez (1995), Maingueneau (1998), Rastier (2001).

A partir de ce qui a été revu dans le modèle de l'ISD, nous avons élaboré un instrument d'analyse sous forme de grille. Dans le Tableau 1 (voir ci-dessous), nous présentons une version condensée de cet instrument². Comme nous pouvons l'observer, la grille reprend le modèle de l'ISD - tel qu'il est présenté dans Bronckart (1997) – et introduit, réorganise et

² Dans l'article Coutinho *et al.* (à paraître), nous présentons une version détaillée de notre instrument d'analyse.

reformule quelques éléments, signalés en gras dans la grille. Remarquons qu’un des aspects centraux (mais ce n’est pas le seul) de notre proposition a trait à la considération d’unités appartenant à d’autres systèmes sémiotiques, qui dépassent le linguistique (à titre d’exemple l’iconique). Ceci se justifie par le fait que nous travaillons sur des textes qui généralement sont plurisémiotiques. Pour cette raison, la relation entre différents systèmes est fondamentale pour appréhender l’organisation du texte.

Du point de vue méthodologique, notre travail a avancé avec l’application de la grille aux textes ; ceci nous a permis d’effectuer un travail allant dans deux directions complémentaires : d’un côté, nous avons identifié des récurrences significatives ou des traits communs dans les exemplaires de textes du même genre ; d’un autre côté, nous avons observé les spécificités de réalisation que ces traits assument dans chaque exemplaire. Cette étape a permis de déterminer quels sont les « paramètres du genre », c’est-à-dire ce que prévoit le genre en question et comment il est actualisé dans les textes. Ceci est visible à travers ce que nous avons dénommé les « mécanismes de réalisation textuelle ».

Ainsi, nous défendons l’idée que les « paramètres » permettent d’observer l’identité d’un genre et les « mécanismes » l’identité d’un texte empirique déterminé.

TEXTE	ARCHITECTURE INTERNE	Infrastructure générale ou plan de texte	Constitution d'unités textuelles (UTs)	Verbales	Typifiées	Types de discours ; Séquences ; Autres
					Autres	Identité thématique ; Identité illocutoire ; Autres
					Non verbales (iconiques, semio-graphiques, autres)	
				Mixtes		
				Modalités d'articulation entre UTs	Articulation entre unités typifiées	
			Articulation entre UTs quelconques			
			Disposition des UTs	Spatio-temporelle		
				Rhétorique		
				Autres		
		Mécanismes de textualisation	Connexion (segmentation/liaison)	Organisateurs textuels	Verbaux	
					Non verbaux	
			Cohésion nominale			
			Cohésion verbale			
		Autres				
		Mécanismes de prise en charge énonciative	Distribution des voix			
Expression des points de vue						
		PARAMÈTRES DU GENRE				
		MÉCANISMES DE RÉALISATION TEXTUELLE				

Tableau 1 - Un instrument d’analyse

3. Un exemple d'analyse générique et textuelle

3.1. Aspects méthodologiques

Pour exemplifier ce qui est en jeu dans notre proposition, nous prendrons un des genres appartenant à l'objet d'étude de notre projet : l'annonce publicitaire. Étant donné que notre objectif est l'analyse de genre, un corpus de textes appartenant à ce genre a été constitué. Chaque texte de ce corpus a été analysé en détail, en utilisant les différentes catégories observables dans la grille proposée. Dans cette première analyse ont surgi des mécanismes spécifiques à chaque texte. Ensuite, notre travail a résidé dans la détection des récurrences et des divergences observables dans les textes. Donc, dans cette phase, l'analyse textuelle est un procédé qui rend possible la différenciation entre phénomènes singuliers et phénomènes récurrents et, simultanément, permet de trouver les traits communs sous-jacents aux phénomènes dans les divers textes.

Il serait impossible d'explorer ici toutes les dimensions semio-linguistiques comprises dans l'organisation textuelle et l'organisation générique. Aussi, nous avons sélectionné seulement un des phénomènes : l'articulation entre unités de différente nature sémiotique, c'est-à-dire, ce qui dans la grille est désigné comme « Articulation entre UTs quelconques ». Comme on peut l'observer dans la grille, les modes d'articulation entre unités textuelles constituent un des aspects qui relèvent de l'infrastructure générale du texte. Le choix de cet aspect pour notre exemplification repose sur deux raisons : d'un côté, ce choix entend mettre en avant la pertinence de considérer les unités textuelles de nature sémiotique non verbale pour l'analyse de textes et de genres qui, comme on l'a dit, est un des éléments que nous avons introduit dans le modèle du ISD ; d'un autre côté, ce phénomène est d'une grande importance dans l'analyse de ce genre en particulier.

L'annonce publicitaire est un genre qui explore de diverses manières les relations possibles entre les systèmes sémiotiques différenciés. De fait, la présence d'unités verbales et non verbales dans les annonces est un trait caractéristique de ce genre – un paramètre, dans notre terminologie – comme cela a déjà été souligné par plusieurs chercheurs (cf. à titre d'exemple, Adam & Bonhomme, 1997 : 55-70).

Le bref commentaire que nous ferons ici cherche à démontrer de quelle manière un des paramètres relatifs à l'articulation entre unités verbales et non verbales est actualisé entre deux unités : d'un côté, une unité non verbale que nous appellerons « image principale », qui peut apparaître éventuellement dans les textes et qui se distingue par la taille et la localisation spatiale ; d'un autre côté, une unité verbale qui intervient aussi de manière variable, que nous dénommerons « appel » ou « titre », constituée d'une (ou plusieurs) phrase(s) mise(s) en évidence (par la typographie et/ou la segmentation).

3.2. Un paramètre, deux mécanismes

De l'analyse de notre corpus, qui réunit une ample variété d'annonces publicitaires portugaises, surgit l'unité « image principale » (photo ou dessin de plus grande taille et/ou mise en évidence), qui, quand elle existe, peut présenter l'objet du discours (produit, endroit, service... divulgué), de manière à ce que ce ne soit pas un simple accessoire ou élément illustratif. Un des rôles (ou une des fonctions) que cette unité non verbale peut avoir est la « convocation » à la lecture de l'annonce et, simultanément, l'introduction (ou instauration) du référent. Pour ce, dans l'unité verbale que nous dénommons « appel » (ou titre), il est possible d'employer des formes référentielles avec valeur anaphorique (pronoms et syntagmes nominaux introduits par un démonstratif). De cette manière, l'unité verbale ne peut être interprétée que par la présence de l'unité non verbale. Ceci démontre l'existence d'une articulation fondamentale dans l'organisation des annonces: une relation de dépendance dans laquelle se construit une chaîne référentielle. Dans cette chaîne, l'unité non

verbale a la fonction d'introduction du référent et l'unité verbale constitue un lien de reprise référentielle, qui pourra, éventuellement, continuer dans la section rédactionnelle dénommée « corps de texte ».

Cette caractérisation constitue un paramètre du genre de l'annonce publicitaire pour ce qui est de l'articulation des différentes unités textuelles (verbales ou non verbales). Ceci signifie que les textes sont construits en prenant en considération cette possibilité d'articulation entre unités. Ce paramètre est actualisé de manière différente dans chaque texte. Pour illustrer notre propos, nous allons aborder les mécanismes de réalisation textuelle qui actualisent ce paramètre à travers deux exemples.



Figure 1 - Annonce de la voiture Alfa 147

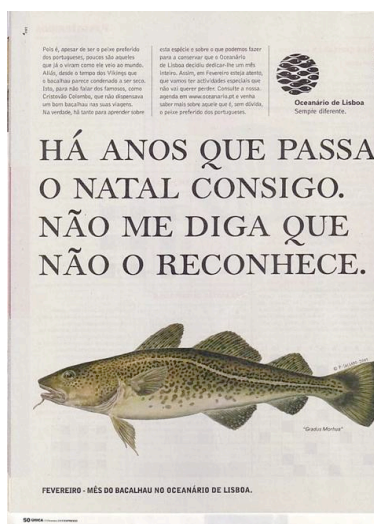


Figure 2 - Annonce de l'Océanarium de Lisbonne

Dans l'exemple 1 (annonce de la voiture Alfa 147), nous voyons l'image de la voiture au centre de la page. L'« appel » (le titre) est distribué en deux sections: a) sur fond rouge, présence des possessifs « seu » /votre/ et « sua » /votre/ ancrés dans l'image ; b) phrase avec typographie plus grande, au centre de la page, sur l'image, emploi du pronom personnel « ele » /il/, référent à l'image.

Dans l'exemple 2 (annonce de l'Océanarium de Lisbonne), l'image principale présente le dessin d'une morue. Dans l'« appel », il y a une anaphore comme sujet du verbe « passa » /passe/ et le pronom « o » /le/ qui fait référence au poisson dessiné. C'est seulement dans le corps du texte (section rédactionnelle) que le terme « bacalhau » /morue/ est explicité.

Comme nous le voyons, si les deux textes actualisent le même paramètre, chacun d'eux emploie des unités différentes (photo vs dessin, possessifs et pronom personnel vs anaphore et pronom personnel) et des organisations spécifiques (par exemple, la présence d'un ou de deux appels, la structure et sa localisation dans la page).

Aspect de l'organisation interne			Paramètre du genre	Mécanisme Exemple 1	Mécanisme Exemple 2
Infrastructure générale ou plan de texte	Modalités d'articulation entre UTs	Articulation entre UTs quelconques	Relation de dépendance entre l'unité non verbale « image principale » et l'unité verbale « appel » dans laquelle se construit une chaîne référentielle. L'unité non verbale a la fonction d'introduction du réfèrent et l'unité verbale constitue un lien de reprise référentielle	L'image de la voiture au centre de la page introduit l'objet de discours et instaure un réfèrent. L' « appel » est distribué en deux sections: a) sur fond rouge, présence des possessifs ancrés dans l'image ; b) phrase avec typographie plus grande, au centre de la page, sur l'image, emploi du pronom personnel « ele »/il/, ancrée dans l'image.	L'image principale au centre de la page présente le dessin d'une morue. Dans l' « appel », l'anaphore comme sujet du verbe « passa » /passe/ et le pronom « o » /le/ qui fait référence au poisson dessiné

Tableau 2 - Un exemple

Nous remarquons à travers cette exemplification que les mécanismes sont spécifiques aux textes – voir la différence entre mécanisme de l'exemple 1 et celui de l'exemple 2 – mais ils actualisent une même caractéristique associée au genre dans lequel les textes s'inscrivent. Si nous incluons plus d'exemplaires de notre corpus, nous observerions que dans chaque texte empirique qui actualise ces paramètres les mécanismes réalisés sont différents. Quand nous effectuons l'analyse textuelle de ces exemplaires, nous trouvons ce qu'ils ont de plus spécifique et singulier ; quand nous réalisons une analyse de genre qui rapproche les deux textes, nous cherchons à caractériser ce qu'il y a de commun dans les mécanismes des différents textes. Dans ces exemples, les deux textes actualisent le trait commun de l'occurrence d'un certain mode d'articulation entre les deux unités textuelles.

4. Conclusion

La proposition que nous vous avons ici sommairement présentée entend montrer que analyse de textes et analyse de genres constituent deux classes d'approches différentes des productions langagières. L'analyse textuelle peut être une fin en soi ou peut être, comme dans notre recherche, un procédé pour entreprendre une analyse de genre.

L'analyse textuelle cherche à décrire les spécificités d'un texte considéré dans son individualité. Dans le processus de production textuelle, tout texte actualise des paramètres du genre dans lequel il s'inscrit par des mécanismes de réalisation textuelle. Cependant, l'actualisation des paramètres n'est pas une simple reproduction, et ce, pour deux raisons : d'un côté, parce que la relation entre paramètres génériques et mécanismes de réalisation textuelle n'est pas biunivoque (par exemple, un même paramètre peut être actualisé par différents mécanismes, comme nous l'avons vu dans les exemples) ; d'un autre côté, dans chaque acte de production textuelle, la configuration paramétrique associée au genre dont le texte relève est adaptée aux propriétés particulières de la situation dans laquelle l'action langagière s'insère. Ainsi, un texte est toujours un objet unique, singulier, non répétable. Lorsque nous entreprenons une analyse textuelle, nous décrivons l'objet dans cette singularité.

L'analyse de genres, quant à elle, cherche à identifier ce qui, pour les textes singuliers fonctionne comme modèle d'orientation ou guide, dans le processus de production et aussi, mais de manière différente, dans le processus de compréhension textuelle. Il s'agit d'un relevé d'aspects ou traits récurrents dans des textes associés à une même catégorie générique. L'objectif d'une analyse de genres est de décrire et d'objectiver ce qui constitue, comme

représentation sociocognitive, le modèle du genre en question pour les producteurs et les récepteurs.

Si nous avançons que les deux approches peuvent être réalisées en appliquant un même instrument d'analyse (représenté schématiquement par une grille dans notre proposition), il faut distinguer deux catégories associées à chaque espèce d'approche. Ainsi, paramètres et mécanismes sont les résultats des données que l'on obtient dans chaque plan d'analyse (générique et textuel).

La distinction théorico-méthodologique entre « analyse textuelle » et « analyse de genres » est pertinente non seulement pour une Théorie Linguistique des Textes, mais aussi pour les domaines de la linguistique appliquée, comme l'enseignement-apprentissage des langues et la traduction. En effet, assumer la différenciation des phénomènes qui sont relatifs au genre de ceux qui sont propres au texte lui-même, en tant qu'objet individuel, constitue un espace qui nécessairement doit être pris en considération dans l'approche des unités de communication linguistique que sont les textes. La raison principale de cette affirmation se trouve dans la manière dont nous concevons la relation entre le *général* et le *particulier* : chacun de ces plans s'implique mutuellement, une fois que les textes sont des exemplaires de genre et les genres sont des formes textuelles relativement stabilisées dans et par l'utilisation du langage. Un travail didactique ou empirique (comme la traduction) avec des textes a tout à gagner si on admet la singularité de l'objet et, parallèlement, son affiliation générique. Cependant, il faut souligner que les instruments élaborés dans le cadre de recherches théoriques, comme la nôtre, ne peuvent être transférés sans aucune autre réflexion vers les autres domaines ; de fait, et spécialement dans le cas de la didactique des langues et de la formation des traducteurs, il faut encore développer des instruments spécifiques qui rendent compte des dimensions pertinentes pour ces domaines. Dans tous les cas, notre proposition donne quelques pistes dans cette voie.

Références

Références bibliographiques

- ADAM, Jean-Michel (1999). *Linguistique textuelle. Des genres de discours aux textes*. Paris : Nathan.
- ADAM, Jean-Michel (2005). *La linguistique textuelle. Introduction à l'analyse textuelle des discours*. Paris: Armand Colin.
- ADAM, Jean-Michel, BONHOMME, Marc (1997). *L'argumentation publicitaire. Rhétorique de l'éloge et de la persuasion*. Paris : Nathan.
- BERNARDEZ, Enrique (1995). *Teoría y epistemología del texto*. Madrid: Cátedra lingüística.
- BRONCKART, Jean-Paul (1997). *Activité langagière, textes et discours. Pour un interactionisme socio-discursif*. Lausanne: Delachaux et Niestlé.
- BRONCKART, Jean-Paul (2004). Les genres de textes et leur contribution au développement psychologique, *Langages*, 153 : 98-108.
- COUTINHO, Maria Antónia, ALVES, Marisa, GONÇALVES, Matilde, MIRANDA, Florencia & PINTO, Rosalice (à paraître), *Parâmetros de géneros e mecanismos de realização textual – aspectos teóricos*.
- MAINGUENEAU, Dominique (1998). *Analyser les textes de communication*. Paris : Dunod.
- RASTIER, François (2001) *Arts et sciences du texte*. Paris : PUF.

Publicités

Texte Alfa 147: Magazine National Geographic (Portugal), Juillet 2004.

Texte Oceanário de Lisboa: Magazine Única, Expresso (Portugal), Février 2004.

RÉGULARITÉS ET IRRÉGULARITÉS DANS LA POLYSÉMIE DES NOMS D'AFFECT : LE CAS DE L'EXPRESSION DE LA SOURCE DE L'AFFECT

Vannina GOOSSENS

vannina.goossens@u-grenoble3.fr

LIDILEM, Université Stendhal Grenoble3

BP 25 - 38040 Grenoble cedex 9

Abstract : The work presented in this article is an exploratory study of the polysemy of affect nouns in which we focus on a particular type of regular polysemy : the cases where the affect noun refers to the source of this affect. We will try to identify the linguistic characteristics of this type of polysemy and to determine whether it gives rise to significant groups of nouns.

1. Introduction

Les noms d'affect (désormais *N_affect*) tels *peur*, *colère*, *joie*, *amour*, etc., sont sémantiquement complexes et, de plus, fortement polysémiques. Nous nous proposons d'étudier les caractéristiques de cette polysémie dans le cadre d'un traitement unitaire des *N_affect* permettant de poursuivre l'entreprise typologique initiée sur ces noms à partir de leurs propriétés combinatoires (Tutin *et al.*, 2006, Goossens, 2005). Nous nous demanderons en particulier si ces noms obéissent à des modèles de polysémie spécifique, et sur quelles bases on peut fonder les régularités observées. Nous émettons en effet l'hypothèse qu'il existe des schémas de polysémie régulière qui pourraient s'avérer caractéristiques de cette classe de noms et susceptibles de nous renseigner sur leurs caractéristiques sémantiques. Afin de tester cette hypothèse, nous avons décidé de nous intéresser dans un premier temps à un seul type de dérivation polysémique régulière, les cas où le *N_affect* semble exprimer en réalité la source de cet affect.

2. Analyser la polysémie des *N_affect*

2.1. Critères de définition des *N_affect*

Les *N_affect* ont donné lieu à de nombreux travaux ces quinze dernières années. Des travaux portant essentiellement sur leurs propriétés distributionnelles (Gross, 1995, Balibar-Mrabti, 1995, etc.) mais aussi combinatoires (Mel'čuk et Wanner, 1996, Tutin *et al.*, 2006, etc.). Ces travaux ont généralement mis en évidence que cette classe de noms est difficile à définir, aussi bien du point de vue de ses caractéristiques sémantiques que de ses caractéristiques syntaxiques et distributionnelles.

Sémantiquement tout d'abord, la frontière entre ce qui est un *N_affect* et ce qui ne l'est pas n'est pas facile à établir. La distinction entre diverses classes de noms psychologiques (sensations, états, qualités, attitudes, affects etc.) se heurte au fait que ces découpages sont intuitifs et langagiers (voir notamment les réflexions de Jean-Claude Anscombre, 1995 et 1996, sur le sujet). Ainsi, la polysémie de certains noms semble notamment faire qu'ils peuvent entrer dans plusieurs catégories en fonction de leurs emplois. Par exemple, la *douleur* peut être une sensation dans *ressentir une douleur vive à l'épaule* ou un affect dans *depuis ce drame Claire est accablée de douleur*). Nous n'entrerons pas ici dans le débat terminologique entre *sentiment*, *émotion*, *affect*... et nous choisirons une étiquette conventionnelle volontairement très large, celle d'affect.

Au niveau de leurs caractéristiques syntaxiques, les critères définitoires classiquement mentionnés dans la littérature comme la compatibilité avec les verbes supports *éprouver* et *ressentir* ou encore avec *un sentiment de* ne fonctionnent pas pour tous les noms que l'on pourrait considérer intuitivement comme renvoyant à un affect (cf. notamment Leeman, 1995). Par exemple : *ressentir de l'émotion* semble peu naturel. Inversement, des noms qui ne renvoient pas à des affects peuvent s'accorder de ces critères (*ressentir de la faim*).

Nous nous trouvons cependant devant la nécessité de donner une définition, même imparfaite, de ce que nous entendons par N_affect, afin de ne pas ouvrir la liste à l'infini. Nous avons donc, dans un premier temps, tenté de concilier critères sémantiques et critères linguistiques en considérant comme tels les noms :

- Renvoyant à un processus psychologique ;
- Compatibles avec les verbes supports *éprouver* et *ressentir* et avec *un sentiment de*
- Ayant une structure actancielle à deux ou trois actants : ils ont obligatoirement un expérienceur, qui ressent l'affect, une cause et/ou un objet à cet affect.

Etant donné les limites que nous avons énoncées précédemment quant à ces critères, nous avons fait le choix de prendre également en compte des noms ne satisfaisant pas forcément les deux derniers critères mais qui peuvent pourtant renvoyer à un affect, au moins dans certains de leurs emplois. C'est par exemple le cas d'un nom comme *orgueil*, qui s'apparente généralement à une qualité (si l'on reprend les catégories traditionnellement mentionnées dans la littérature) mais qui peut également être considéré comme un affect, comme par exemple dans *ressentir une grande bouffée d'orgueil*. Comme nous allons le voir par la suite, ce choix a des impacts importants lorsque l'on s'intéresse à la polysémie. Notre hypothèse, en effet, est que l'étude de la polysémie des N_affect pourrait nous amener à une meilleure définition de cette classe, ou en tout cas à mieux la cerner à travers les dérivations sémantiques auxquelles elle donne lieu et à mettre en évidence qu'il existe plusieurs sous-types de N_affect.

2.2. Les types de polysémie

Il existe deux types de polysémie, qui ont notamment été étudiées par Apresjan (1974), Cruse (1986), Pustejovsky (1991, 1995), Nunberg & Zaenen (1997), etc. : une polysémie régulière ou systématique et une polysémie que l'on pourrait appeler irrégulière. La polysémie régulière ou systématique consiste en une dérivation sémantique existant pour tous les éléments d'une classe sémantique, ou en tout cas pour une grande partie d'entre eux, alors que l'on parlera de polysémie irrégulière lorsque cette dérivation sémantique est propre à un élément particulier. Si on prend l'exemple de noms d'administrations publiques comme *école* ou *hôpital*, pour sortir du champ des affects, on remarque qu'*école* peut avoir le sens *d'école de pensée*, mais pas *hôpital* : c'est un cas de polysémie irrégulière. Ces noms donnent également lieu à de la polysémie régulière. En effet, outre le sens d'administration (*l'hôpital public/l'école publique*), il peuvent tous deux véhiculer le sens de 'bâtiment abritant cette administration' (*l'hôpital/l'école est en ruine*) ou de 'personnel employé par cette administration' (*l'hôpital/l'école est en grève*). Ces dérivations sémantiques, ou schémas de polysémie régulière, nous renseignent sur les caractéristiques des noms concernés, en l'occurrence une administration occupe des locaux, emploie du personnel, etc.

Les N_affect ne dérogent pas à cette règle. D'une part, ils donnent évidemment lieu à une multitude d'acceptions polysémiques propres à un nom en particulier. Par exemple, *horreur* peut désigner un affect proche de la peur (*avoir une vision d'horreur*), ou alors un affect proche du dégoût (*les épinards lui font horreur*). *Bonheur* peut, quand à lui, désigner évidemment un affect ou alors avoir le sens non affectif de 'chance' (*porter bonheur*). Et ainsi de suite, pour les différents éléments de la classe.

D'autre part, ils donnent également lieu à des acceptions polysémiques régulières, et c'est ce qui nous intéresse plus particulièrement ici. Voici quelques régularités que l'on peut observer pour la classe des N_affect.

Dans *Pierre est d'une grande gaieté*, le N_affect a le sens d'une 'qualité de la personne', une qualité permanente, c'est-à-dire la propension à ressentir si souvent cet affect que cela en vient à le qualifier. On peut ainsi avoir *être d'une grande tristesse*, *d'une grande jalousie* etc. Le N_affect peut également prendre le sens d'une 'qualité causant l'affect', comme dans *un paysage d'une grande tristesse*. Le paysage a ici la qualité de provoquer cet affect. Le nom peut également prendre le sens de 'source' de l'affect comme dans *Paul est la joie de sa mère*, où *joie* peut être paraphrasé par *source de la joie*. Autre exemple, dans *Paul fait des tendresses à Claire*, le nom a ici le sens d'actes qui dénotent la présence de cet affect chez celui qui les effectue. Enfin, dans *Claire a des amitiés haut placées*, *amitié* ne renvoie pas à l'affect en lui-même mais à des 'relations' avec des personnes qui impliquent cet affect.

Lorsqu'on observe ces différents exemples, on pressent que les N_affect n'ont pas un comportement homogène vis-à-vis de ces dérivations sémantiques. Par exemple, *gaieté* semble pouvoir exprimer la 'qualité d'une personne' (*Paul est d'une grande gaieté*), une 'qualité causative' (*un paysage d'une grande gaieté*) mais pas la 'source' de cet affect (*Paul est la gaieté de sa mère*), ni un 'acte' (*faire des gaietés*). Il s'agira donc à la fois de mettre en évidence quels N_affect permettent quel(s) type(s) de polysémie, d'essayer d'en analyser les raisons et enfin de déterminer les caractéristiques linguistiques de ces divers types de polysémie. Afin de vérifier si ces différents objectifs sont réalisables, du moins en partie, nous allons nous intéresser ici à la polysémie de type 'source'.

3. Méthodologie

Cette étude exploratoire est basée sur un corpus composé de 10 textes du XIX^{ème} siècle collectés dans le cadre d'un projet¹ qui porte sur l'étude des marqueurs linguistiques de la subjectivité. Ce corpus représente 830 000 mots et est essentiellement composé de romans².

Nous avons également utilisé un corpus complémentaire, afin de vérifier la fréquence de certaines structures mais aussi pour bénéficier de textes plus contemporains. Il est issu de la base catégorisée de Frantext et composé de romans qui couvrent la période 1990-2005, soit 22 textes et plus de 1,5 millions de mots.

Nous avons recherché dans ces textes une importante liste de N_affect : plus de 250. Ces noms correspondent à la définition que nous avons proposée précédemment, définition dont le caractère assez large nous permet d'essayer d'appréhender la polysémie des N_affect de la manière la plus exhaustive possible. La prise en compte de noms s'apparentant la plupart du temps à des qualités (*orgueil*, *bonté*...), à des états (*solitude*, *béatitude*...) ou encore à des attitudes (*agressivité*, *tendresse*...) nécessite cependant quelques précautions d'ordre méthodologique. En effet, le sens de la dérivation sémantique s'en trouve affecté. Prenons l'exemple de l'acception 'qualité d'une personne' évoquée dans la partie précédente. Pour un nom comme *tristesse* la dérivation sémantique va s'effectuer dans le sens 'affect' vers 'qualité'. A l'inverse, pour un nom comme *orgueil*, le sens de la dérivation sera orienté de 'qualité' vers 'affect'. Cela a donc notamment des répercussions sur les constructions syntaxiques de ces différents noms.

¹ Plan pluri-formation piloté par le laboratoire LIDILEM (Université Stendhal Grenoble 3).

² *Les contes*, de Perrault, *Le petit chose*, *Les contes du lundi*, *Les lettres de mon moulin*, d'Alphonse Daudet, *La petite fadette*, de Georges Sand, *Le droit à la paresse*, de Paul Lafargue, *L'île mystérieuse*, de Jules Verne, *Une vie*, de Maupassant, *Colomba*, de Prosper Mérimée et *Le mystère de la chambre jaune*, de Gaston Leroux.

Nous avons tenté d'identifier dans notre corpus les cas renvoyant à une polysémie de type 'source', de déterminer ses caractéristiques linguistiques et les regroupements qu'elle opère parmi les N_affect.

Nous laissons de côté pour l'instant la question des fréquences. Dans cette étude exploratoire, le corpus n'est utilisé que pour repérer des exemples attestant les différents types de polysémie. Ces phénomènes sont en effet peu productifs et nécessitent l'usage de très gros corpus pour pouvoir envisager un traitement statistique des données.

4. Analyse des résultats

La dérivation polysémique régulière de type 'source' que nous avons précédemment illustrée par la phrase *Paul est la joie de sa mère* assimile la 'source' d'un affect à cet affect en lui-même : *Paul = la joie de sa mère*. Nous noterons désormais N_source_affect les noms qui véhiculent cette acception dérivée par la polysémie régulière de type 'source'.

En fonction de la structure actancielle du nom, cette source peut être sur le plan sémantique, soit une cause, comme pour *chagrin* (*le chagrin de X devant Y*) en (1), soit un objet, comme pour *amour* (*l'amour de X pour Y*) en (2).

(1) *la pluie [...] était le premier gros **chagrin** de son existence* (Maupassant, Une vie)

(2) *Tu étais ma vie, mon rêve, mon seul espoir, mon seul **amour*** (Maupassant, Une vie)

Nous allons dans un premier temps tenter de définir les caractéristiques des constructions comportant des N_source_affect.

4.1. Les constructions comportant des N_source_affect

L'observation des N_source_affect présents dans notre corpus a mis en évidence que ce type de dérivation polysémique était généralement véhiculé par le biais de deux constructions syntaxiques privilégiées.

Premièrement, une construction du type *ma fille est ma joie* où la source de l'affect est mise en relation avec le N_source_affect à l'aide d'un verbe d'état (dans l'immense majorité des cas il s'agira du verbe *être*, mais il semble que le verbe *devenir*, et peut-être d'autres encore peuvent également fonctionner) comme dans les trois exemples suivants :

(3) *je suis une **peine** et un **embarras** à ma famille* (Sand, La petite Fadette)

(4) *Ce père Cornet. Ç'a été le satan de mon enfance, ma **passion** douloureuse* (Daudet, Les contes du lundi)

(5) *Le "Bonjour, Boucoyran!" était devenu un **supplice** pour moi* (Daudet, Le petit chose)

La seconde structure syntaxique dans laquelle entrent les N_source_affect est une construction détachée, comme dans les deux exemples suivants :

(6) *sous sa rude moustache, **effroi** des chiens et des traîneurs de banc, se cachait un bon sourire attendri* (Daudet, Les contes du lundi)

(7) *Quinze ans, de gros pieds, de gros yeux, [...] et l'allure d'un valet de ferme, tel était le marquis de Boucoyran, **terreur** de la cour des moyens et seul échantillon de la noblesse cévenole au collège de Sarlande* (Daudet, Le petit chose)

Enfin, on relève un certain nombre d'autres occurrences où l'on est bien en présence d'un N_source_affect mais où l'on n'identifie pas de structure syntaxique aussi contrainte. Par exemple :

(8) *Eux qui trouvent les journées si longues, et qui sont si heureux quand c'est leur tour d'aller à la terre... Dans la belle saison, ce grand **bonheur** leur arrive tous les mois* (Daudet, Les lettres de mon moulin)

(9) *Or ; savez-vous quel est ce quelqu'un de la ville que le petit Chose veut voir avant de partir ? C'est la fabrique, cette fabrique qu'il aimait tant et qu'il a tant pleurée !... c'est le jardin, les ateliers, les grands platanes, tous les amis de son enfance, toutes ses **joies** du premier jour...* (Daudet, Le petit Chose)

Une étude sur un corpus plus étendu pourrait peut-être permettre de dégager d'autres régularités syntaxiques.

Quelle que soit la structure syntaxique de l'énoncé dans lequel apparaît le N_source_affect, celui-ci est toujours accompagné d'un élément (de statut syntaxique différent en fonction de cette structure) renvoyant à la source de l'affect, comme par exemple *la pluie* en (1). Dans l'optique de mieux caractériser ce type de dérivation polysémique, nous allons maintenant tenter de voir si l'on peut relever des restrictions du côté de l'élément source, puis du N_source_affect lui-même.

4.2. La structure sémantique de la bipartition source-N_source_affect

4.2.1. Caractéristiques de l'élément source

A partir des exemples relevés dans notre corpus, il semble difficile d'identifier des restrictions sémantiques de ce côté-là étant donné la diversité des éléments relevés. On constate que la source de l'affect peut être :

- Une personne comme dans les exemples (2), (3), (4) et (7)
- Une partie du corps comme dans l'exemple (6)
- Un objet concret comme dans l'exemple suivant :

(10) *c'est l'élixir du père Gaucher, la **joie** et la santé de notre Provence, me fit le brave homme d'un air triomphant* (Daudet, Les lettres de mon moulin)

- Une qualité, comme dans l'exemple suivant :

(11) *Cette facilité de donner était du reste un des grands **bonheurs** de leur vie* (Maupassant, Une vie)

- Une activité, comme dans l'exemple ci-dessous :

(12) *Il ne reculerait devant aucune tâche, et, chasseur passionné, il ferait un métier de ce qui, jusqu'alors, n'avait été pour lui qu'un **plaisir*** (Verne, L'île mystérieuse)

- Un événement, comme dans les exemples (1) et (8), etc.

Nous constatons également une grande diversité du point de vue de leurs réalisations syntaxiques. Une caractérisation plus précise, aussi bien sémantique que syntaxique, de cet élément pourrait permettre d'affiner la caractérisation des structures syntaxiques présentées précédemment. En effet, essentiellement dans le cas de la construction détachée, la structure n'est pas suffisamment contrainte pour pouvoir, par exemple, envisager un traitement automatique ou semi-automatique de ce type de polysémie.

Nous allons maintenant tenter de voir quels noms, parmi ceux de notre liste de N_affect, sont susceptibles d'être des N_source_affect.

4.2.2. Caractéristiques des N_source_affect

Comme toutes les dérivations polysémiques régulières que nous avons mentionnées ('qualité', 'qualité causative', 'relation', etc.) la dérivation de type 'source' ne fonctionne pas avec tous les N_affect. Ainsi, *amour, bonheur, joie, effroi, passion, chagrin, embarras, orgueil, peine, remords* ou encore *plaisir* peuvent être des N_source_affect, mais cela ne semble pas être le cas de *gaieté, mépris, respect, estime, mélancolie, bonté, pudeur*³ etc. Dans un premier temps, nous avons souhaité vérifier si nous pouvions recouper ces observations avec la distinction, effectuée notamment par Anscombe (1995) puis par Mathieu (1999), entre :

- Noms exogènes, ou hétéro-initiés, c'est-à-dire des noms qui renvoient à des affects qui ont une origine extérieure à l'expérimenteur, qui sont en quelque sorte des réactions à un stimulus (ex : *peur, tristesse, surprise*).
- Et noms endogènes, ou auto-initiés, qui renvoient à des affects qui prennent leur origine en l'expérimenteur lui-même du fait d'un jugement que celui-ci porte sur un fait extérieur (ex : *mépris, respect, admiration*).

Cette distinction est fondée sur un certain nombre de critères formels. Par exemple, les noms endogènes se construisent avec les prépositions *pour* et/ou *envers* (*le mépris de Pierre pour/envers ses semblables, l'amour de Paul pour Claire, etc.*) alors que les exogènes se construisent avec *devant* et/ou *à la vue de* : *la joie de Max à la vue de sa fille, le désespoir de Paul devant sa copie, etc.*

Si l'on regarde la structure actancielle des noms qui entrent dans ces deux catégories, on observe que les exogènes regroupent les N_affect biactanciels, comme *joie* ou *chagrin* (*joie/chagrin de X devant Y*), dont le deuxième actant (Y) est une cause. Les endogènes, quant à eux, regroupent les noms bi-actanciels avec un second actant objet (*amour de X pour Y, passion de X envers Y*) et les noms tri-actanciels, comme *mépris, respect* ou *estime*, qui ont à la fois un actant objet (Y) et un actant cause (Z) (*mépris/respect/estime de X envers Y devant Z*).

Pour ce qui est de la dérivation polysémique de type 'source', on peut émettre l'hypothèse que les N_source_affect doivent être préférentiellement des noms exogènes, dans le sens où ils désignent clairement un élément extérieur à l'expérimenteur comme étant sa source. Cela semble effectivement être le cas, la quasi-totalité des N_source_affect de notre corpus étant des noms exogènes (*joie, bonheur, peine, etc.*). Deux remarques à présent : premièrement, nous avons relevé deux N_source_affect endogènes, *amour* et *passion* ; et deuxièmement, tous les noms exogènes ne peuvent pas être N_source_affect (par exemple cela ne semble pas être le cas de *gaieté* ou *mélancolie*).

Pour ce qui est de *amour* et *passion*, il faut tout d'abord préciser qu'ils ont un comportement très spécifique, plus particulièrement *amour* qui mériterait un traitement particulier. Ils peuvent notamment tous deux véhiculer ce sens un peu concret de 'source' en dehors de toute structure syntaxique particulière : *un amour* est quelqu'un envers qui on pourrait ressentir de l'amour, *une passion* est une activité pour laquelle on ressent de la passion. Cependant, nous pouvons remarquer qu'ils font partie des rares endogènes qui n'ont que deux actants (un expérimenteur et un objet), la plupart des noms de cette classe étant tri-actanciels, c'est-à-dire dirigés vers un objet pour une cause particulière. On peut ainsi avoir un exemple du type *le mépris de Paul pour Pierre devant la lâcheté de son geste* mais difficilement *l'amour de Paul pour Marie à cause de sa gentillesse/de sa beauté*. La

³ Il est très difficile de dire quels sont les noms qui acceptent ou qui n'acceptent pas ce type de polysémie sans s'appuyer sur une étude de corpus, c'est pourquoi nous restons prudente concernant les jugements d'acceptabilité.

dérivation de type 'source' ne sélectionnerait peut-être ainsi que des noms bi-actanciels, que le deuxième actant soit une cause ou un objet.

Enfin, le fait que tous les noms exogènes ne puissent pas être des N_source_affect nous amène à proposer la distinction suivante, entre les noms :

- Qui ont une source identifiable : *joie, bonheur, chagrin, amour etc.*
- Et ceux qui n'ont pas toujours de source clairement identifiable : *tristesse, gaieté, mélancolie, etc.*

Cette dernière distinction renvoie à la façon dont nous pouvons concevoir les affects, avec d'un côté des affects qui sont toujours ressentis pour une raison précise et connue de l'expérienceur, et d'un autre, ceux qui peuvent être ressentis sans que celui-ci ait toujours réellement conscience de celle-ci (voir par exemple la typologie proposée par Johnson-Laird et Oatley, 1989). Cela nous permettrait ainsi de distinguer *joie* et *gaieté*, ou alors *chagrin* et *tristesse*, les premiers étant généralement liés à un événement précis alors que les seconds peuvent être ressentis parfois plus ou moins sans raison objective ou consciente.

La dérivation régulière de type 'source' semble ainsi sélectionner préférentiellement des noms qui ont une cause précise (*joie, chagrin, effroi*) ou bien des noms qui sont uniquement dirigés vers un objet, sans qu'il n'y ait de cause particulière à l'existence de cet affect (*amour et passion*).

5. Conclusion

Cette étude exploratoire de la polysémie des N_affect a permis de vérifier en partie nos hypothèses. En premier lieu, nous avons vu que l'on peut dégager des dérivations polysémiques régulières au sein de la classe des N_affect. Dans un deuxième temps, l'étude d'un type particulier de dérivation polysémique régulière, la dérivation de type 'source', nous a amenée à deux constats.

Tout d'abord, sur le plan syntaxique, il est possible de dégager des régularités dans les constructions des N_source_affect. Nous avons ainsi mis en évidence deux structures propres à ces noms : une construction attributive du type *ma fille est ma joie* et une construction détachée du type *Claire, joie de ses parents....* Il y a bien évidemment un certain nombre d'occurrences de N_source_affect pour lesquelles nous n'avons pas pu dégager de construction spécifique.

Sur le plan sémantique, l'étude de la dérivation polysémique de type 'source' a également mis en évidence le fait que tous les N_affect ne semblent pas pouvoir être des N_source_affect. L'observation de la bipartition créée par ce type de polysémie nous a amenée à envisager l'existence de deux classes de noms : d'un côté les noms qui peuvent être N_source_affect qui auraient généralement une source clairement identifiable, que ce soit une cause (*joie, effroi, bonheur, chagrin*) ou un objet (*amour, passion*) ; de l'autre des noms qui ne peuvent pas être N_source_affect et dont l'apparition peut être liée à une cause beaucoup plus diffuse (*gaieté, tristesse, mélancolie*) ou à la fois à une cause et un objet (*mépris, respect, estime*). Cette distinction devra être affinée par la suite, notamment en étudiant les autres types de dérivations régulières (par exemple, les dérivations 'qualité' et 'qualité causative' semblent plutôt sélectionner des noms qui ne semblent pas être N_source_affect comme *gaieté, mélancolie, tristesse...*) mais également à l'aide d'autres critères (il serait intéressant de confronter ces résultats avec les typologies basées sur les collocations que j'ai évoquées précédemment) car elle regroupe des noms de natures différentes notamment de par leur structure actancielle.

L'étude des dérivations polysémiques régulières semble ainsi nous apporter des précisions sémantiques sur les noms qui les acceptent (et donc également sur ceux qui ne les acceptent

pas) et pourrait ainsi nous aider à mieux cerner la classe des N_{affect} comme nous en avons émis l'hypothèse.

Références

- ANSCOMBRE, Jean-Claude (1995). Morphologie et représentation événementielle : le cas des noms de sentiment et d'attitude, *Langue française*, n° 105 : 40-54.
- ANSCOMBRE, Jean-Claude (1996). Noms de sentiment, noms d'attitude et noms abstraits, in *Les noms abstraits, histoire et théories*, Flaux, Nelly, Glatigny, Michel & Samain, Didier (Eds). Lille : Presses Universitaires du Septentrion, 257-273.
- APRESJAN, Jurij D. (1974). Regular polysemy, *Linguistics*, n°142 : 5-32.
- BALIBAR-MRABTI, Antoinette (1995). Une étude de la combinatoire des noms de sentiment dans une grammaire locale, *Langue française*, n° 105 : 88-97.
- CRUSE, D. Alan (1986). *Lexical Semantics*. Cambridge : University Press.
- GOOSSENS, Vannina (2005). Les noms de sentiment : esquisse de typologie sémantique fondée sur les collocations verbales, *Lidil*, n°32 : 103-121.
- GROSS, Maurice (1995). Une grammaire locale de l'expression des sentiments, *Langue française*, n° 105 : 70-87.
- JOHNSON-LAIRD, Philip & OATLEY, Keith (1989). The Language of Emotions : An Analysis of a Semantic Field, *Cognition and Emotion*, vol 2, n° 3: 81-123.
- LEEMAN, Danièle (1995). Pourquoi peut-on dire *Max est en colère* mais non **Max est en peur*, *Langue française*, n° 105 : 55-69.
- MATHIEU, Yvette Yannick (1999). Les prédicats de sentiment, *Langages*, n°136 : 41-52.
- MEL'CUK, Igor & WANNER, Léo (1996). Lexical Functions and Lexical Inheritance for Emotions Lexemes in German, in *Lexical Functions in Lexicography and Natural Language Processing*, Wanner, Léo (Ed), Amsterdam: John Benjamins, 209-278.
- NUNBERG, Geoffrey & ZAENEN, Annie (1997). La polysémie systématique dans la description lexicale, *Langue Française*, n°113 : 12-23.
- TUTIN, Agnès, NOVAKOVA, Iva, GROSSMANN, Francis & CAVALLA, Cristelle (2006). Esquisse de typologie des noms d'affect à partir de leurs propriétés combinatoires, *Langue Française*, n°150 : 32-49.
- PUSTEJOVSKY, James (1991). The Generative Lexicon, *Computational Linguistics*, 17, n°4 : 409-441.
- PUSTEJOVSKY, James (1995). *The Generative Lexicon*. Cambridge : MIT Press.

LA THÉORIE DES MATRICES ET DES ÉTYMONS EN BERBÈRE OU NOUVELLE HIÉRARCHISATION DU LEXIQUE

Cécile LUX

Cecile.Lux@univ-lyon2.fr

Laboratoire Dynamique Du Langage

Institut des Sciences de l'Homme

14, avenue Berthelot

69007 Lyon

Abstract : It is traditionally assumed that the Berber lexis conforms to a root-scheme structure. Our aim is to show that it is possible to go beyond the notion of root, and that Berber words are built on deeper phonetico-semantic relations, on units which have much in common with the etymons that Bohas's TME theory identified for Arabic.

Le berbère est un ensemble de langues parlées sur une zone couvrant tout le nord de l'Afrique, allant de Siwa, en Égypte, jusqu'à l'Atlantique ; de la Méditerranée jusqu'au sud du fleuve Niger. Au niveau historique, nous n'avons aucune trace de l'arrivée des peuples berbères sur ce territoire, ni dans la littérature, ni dans la tradition orale : il s'agirait ainsi du peuple présent en Afrique du nord avant les invasions puniques, romaines et arabes, pour ne citer que celles-ci. Ces différentes dominations successives ont poussé les Berbères à se retrancher dans des zones constituant des protections naturelles : montagnes et déserts. Ainsi, le territoire berbère se présente aujourd'hui sous la forme d'îlots, séparés par des zones où d'autres langues ont pris le dessus. Ces faits historiques permettent d'expliquer l'unité linguistique et culturelle profonde dans le domaine berbère, mais aussi les divergences, engendrées par l'éloignement. Il est d'usage, pour décrire le monde berbère, de parler d'unité dans la diversité.

1. L'organisation lexicale traditionnelle en berbère : schèmes et racines

Au niveau linguistique, les différentes variétés du berbère constituent l'une des six sous-branches de la macro-famille des langues afro-asiatiques, au même titre que le sémitique, dont l'arabe ou l'hébreu font partie. Or, l'un des traits morphologiques communs aux langues afro-asiatiques est que les unités lexicales sont « le produit du croisement d'un schème et d'une racine » (Cohen, 1988 : 11).

L'élément appelé « racine » dans la tradition linguistique sémitique (terme maintenant utilisé dans presque tout le domaine afro-asiatique) rejoint une définition générale de la racine, proposée par Meillet pour l'indo-européen : « [on dit qu']un mot 'appartient' à une racine [quand] il fait partie d'un ensemble de mots ayant en commun un groupe de phonèmes auquel est associé un certain sens général »¹.

Prenons comme exemple un ensemble dérivatif tachelhit, langue berbère parlée dans le sud marocain :

¹ Citation donnée par Cohen dans son article *Racines* (1993 : 165), reprise ici.

zdyγ « habiter + accompli » ; **zyddyγ** « habiter + inaccompli » ; **zdiγ** « habiter + accompli négatif » ; **tazduyt** « le fait d’habiter » ; **amyzday** « habitant »².

Les trois consonnes /zdy/ sont présentes dans tout le paradigme, en corrélation avec le sens général « d’habitat », ce qui correspond à la définition de la racine énoncée ci-dessus. Toutefois, ce qui est spécifique aux racines des langues afro-asiatiques est qu’elles sont exclusivement consonantiques. Aussi, la racine est souvent définie, pour le berbère, comme une consonne ou une succession de consonnes ordonnées « véhiculant une notion générale à l’état brut » (Cadi, 1987 : 37).

Une autre particularité afro-asiatique pour la construction du lexique est la notion de schème : les consonnes radicales, qui ne peuvent exister seules, s’insèrent dans des moules grammaticaux composés de consonnes et de voyelles spécifiques, ainsi que de places vides, qui seront occupées par les consonnes de la racine. Le fait que le nombre de schèmes existant dans ces langues soit assez restreint permet d’identifier la racine avec certitude. Si l’on compare les termes **tazduyt** « le fait d’habiter » et **tamyurt** « le fait d’être habitué », en tachelhit toujours, on peut identifier un schème ta--u-t, composé de deux consonnes et de deux voyelles, dans lequel peuvent être insérées les consonnes de différentes racines, porteuses d’une notion sémantique générale.

Ces deux unités, racine et schème, sont considérées comme les unités de base du lexique berbère, unités qu’on ne peut dépasser.

2. « Un autre type d’organisation du lexique »

2.1. Racines à troisième radicale alternante

Si les notions de racine et de schème semblent assez bien fonctionner pour la description des variétés du berbère, un raisonnement exactement similaire permet toutefois de dépasser ces niveaux.

Prenons comme premier exemple un phénomène bien connu en berbère : les racines à troisième radicale alternante. En effet, il existe, dans ces langues, des séries de termes ayant des racines différentes, mais dont seule la troisième et/ou la quatrième radicale varient, les deux premières étant clables. Ces termes partagent en outre une notion sémantique commune, qui apparaît clairement.

Voyons l’exemple le plus connu, tiré de l’introduction au dictionnaire de Taïfi (1981) :

Racine	Terme tamazight	Sens
FRK	fɾyk	« être ébréché »
FRKK	fɾɾkyk	« se craqueler, se fissurer »
FRRX	fɾurx	« éclore, sortir de l’oeuf »
FRS	fɾurs	« se gercer »
FRFX	fɾɾfɾx	« concasser, casser, briser »
FRRY	fɾury	« s’émietter »

Tableau 1 - Termes à 3^e radicale variante en tamazight³.

Il semble évident que tous ces termes puissent être regroupés autour de la notion générale de « cassure, séparation ». Ainsi, si l’on suit le même raisonnement que celui adopté pour

² Pour les verbes, les morphèmes de personnes et la particule propre à l’inaccompli sont laissés de côté, par commodité. Les données ici sont tirées du dictionnaire de Destaing (1920). Nous avons adopté, pour toutes les données, l’API.

³ Le tamazight est une variété du berbère parlée dans la zone centrale du Maroc.

définir la racine, il convient d'associer ce sens commun aux deux radicales FR, et non à la racine dans son intégralité.

La notion de racine, considérée comme une unité lexicale de base, peut donc être dépassée, et il semble exister des groupes de deux consonnes liées à un sens plus général que celui de la racine. Si l'existence de telles unités se révèle être effective, cela permettrait de mettre en évidence des relations phonético-sémantiques profondes du lexique berbère. Il faut noter que la notion de racine est inopérante ici puisqu'elle n'explique nullement la relation entre les termes cités précédemment, elle les sépare même.

2.2. Racines triconsonantiques ou bases biconsonantiques ?

Dans la littérature, la position par rapport à l'existence de telles bases biconsonantiques, reflet de structures profondes du lexique, est assez ambiguë. En effet, la notion de racine est très largement utilisée dans le domaine berbère, c'est notamment le fondement pour la réalisation des dictionnaires. Or, si de nombreuses discussions portent sur l'origine biconsonantique ou triconsonantique des racines, l'idée la plus répandue est que le nombre de radicales normal est de trois (Galand, 1988 : 72), la présence de nombreuses racines bilitères pouvant être expliquée, le plus souvent, par l'érosion phonétique forte en berbère. Toutefois, plusieurs auteurs ont eu l'intuition de l'existence de groupes biconsonantiques tels que FR, liés à une notion sémantique plus abstraite que celle véhiculée par la racine. Le plus précis allant dans ce sens est justement Galand (1988 : 76) qui, après avoir abordé la question des racines, parle d'un « autre type d'organisation du lexique », signalant l'existence de « bases bilitères », à relier à une « valeur générale ». Plusieurs articles ont aussi été dédiés, nous le verrons, à l'étude d'autres bases biconsonantiques, de manière assez anecdotique et hors de tout cadre théorique.

Ces structures profondes du lexique, structures biconsonantiques véhiculant un sens propre sont donc relativement reconnues en berbère, mais les auteurs ont toujours écarté la possibilité d'une étude poussée et exhaustive de ces unités, la jugeant trop aléatoire pour des données trop réduites en berbère. L'étude de tels groupes reste donc à un stade primaire ; leur réelle nature et leur importance au sein de l'organisation du lexique reste encore inconnue. C'est dans le but d'apporter une contribution à l'identification et à la définition de ces groupes biconsonantiques que j'ai choisi d'étudier ici plus précisément deux de ces unités profondes du lexique, afin de déterminer s'il s'agit d'un phénomène périphérique ou central dans le lexique berbère.

3. Exemples de bases biconsonantiques

3.1. FR et la « séparation », complément d'étude

Nous avons vu précédemment quelques termes, souvent cités, contenant les deux phonèmes /f/ et /r/, clairement reliés à la notion de séparation. Grâce à une étude poussée du lexique tachelhit-français (Destaing, 1920) et de données personnelles recueillies au Maroc, en tachelhit, j'ai recensé, en plus des quelques termes proposés par Taïfi (1981, XI) pour le tamazight, environ 35 termes reliés à cette base FR, véhiculant tous l'idée de séparation. Une liste exhaustive serait un peu fastidieuse pour le lecteur, mais il faut tout de même un certain nombre de termes pour pouvoir apprécier le phénomène. Je donnerai ici les plus significatifs, citant d'abord la racine des lexèmes, puis les termes appartenant à cette racine, transcrits en API, et enfin le sens en français.

Base consonantique FR ⇔ / Séparation /

Sens 0 : « Séparer, établir une séparation dans l'espace »

FRG	fryg	« faire un enclos, enclore » ;
	afrag	« enclos » ; ifrig « barrière d'une maison, d'un jardin, faite de piquets, branches ou broussailles ».

Sens 1 : « Séparer + / cassure, brutalité / »

FRS	fyrsi	« être fendu » ; sfyrsi « fendre » ; afyrsu « éclat de bois »
FRM	frym	« être ébréché » ;
FRY	afyrriu	« brèche, fissure » ;
FRK	sfrukk ^w i	« briser, partager » ;
	afrukk ^w i	« éclat, partie éclatée » ; fyrkki « pelure » ;
FRKγ	førkw ^w γ	« rompre le pain » ;
	afrukkvγ	« brèche faite dans un mur de haut en bas » ; ifyrkγ « cosse » ;
FRR	fruri	« être émietté, égrené » ;
FRZ	afyrzžu	« éclat, coquille » ;
FRSL	fyrsyl	« être gercé » ;
	afursyl	« gerçure ».

Sens 2 : « Séparer avec l'idée de tri, action volontaire et calculée »

• « Séparer un élément externe d'un élément central »

FRK	tifyrkit	« écorce, croûte » ;
FRS	frys	« être tranchant » ;
FRR	tafirt	« crème (formée en se séparant du lait) » ;
•	« Séparer le propre du sale »	
FRN	fyrn	« être nettoyé » ;
FRD	fyrđ	« nettoyer, débayer, curer » ;
FRS	fyrš	« curer, boucharder » ;

Ainsi, on voit que les termes contenant les deux phonèmes FR, ayant quelque rapport avec la notion de séparation, sont nombreux, et que le rapprochement entre forme et sens est inévitable.

Au niveau sémantique, j'ai proposé un classement pour cette base car, si tous les termes se rapprochent de la notion de séparation, des sens distincts peuvent être repérés.

Ce qui est intéressant, ici, est que l'on puisse observer la manière dont une notion sémantique « sous-jacente » prend vie dans le lexique. En effet, pour le sens 0 « Séparer, établir une séparation dans l'espace », la notion de séparation apparaît seule, c'est le sens « de surface » le plus proche de celui proposé en profondeur. En revanche, pour les sens 1 et 2, un sème supplémentaire est ajouté au sens de base, ce qui le fait légèrement varier. Ainsi, le sens de séparation associé à la base biconsonantique FR, notion sémantique très générale sous-jacente, prend vie dans le lexique en étant spécifiée par différents sèmes.

3.2. La base GM, « étendue d'eau plane »

Un complément d'information concernant la nature et les propriétés de ces bases biconsonantiques peut être obtenu grâce à l'analyse d'un groupe de termes ayant des racines différentes, mais reliés pourtant par deux radicaux semblables, GM, et par une notion sémantique générale commune, la notion « d'étendue d'eau plane ». J'ai identifié cette base biconsonantique, détaillée ci-dessous, me plaçant cette fois dans une perspective interdialectale : sont prises en compte les données dont nous disposons pour le tarifit

Base consonantique GM ⇔ / Étendue d'eau plane /

Sens 0 : « Étendue d'eau plane »

GMM	gummu	« masse d'eau animée d'un courant tourbillonnant » -tarifit- ;
GRMM	gyrmyṃ	« former une nappe, stagner » -tarifit- ;
	agyṛəm	« le fait de stagner » -tarifit- ;
	agyṛmam	« lac, étang, cours d'eau » -tarifit- ;
GLMM	gulmim	« ville du sud marocain (station thermale) » -tarifit- ;
	agyṛlmam	« réservoir d'eau naturel, lac, mare » -tamacheq- ;
GRMN	tagyṛmant	« petit abcès, petit lac ».

Sens 1 : « Rive d'un cours d'eau, limite, frontière »

GMḍ/GMḍ	ag ^w mmaḍ	« côté d'une rivière, rive » -tachelhit- ; ag ^w mmaḍ « versant, côté opposé » -kabyle- agymmaḍ « rive d'un cours d'eau, versant, bord » -tarifit- ;
GMD	g ^w mmṽd	« sortir de son lit, franchir, changer de rive » -kabyle- ;
GDM	agyddim	« rive d'un cours d'eau, talus, bord » -tarifit- ;
	ygdyṃ	« entourer de tous côtés ou d'un seul » -tamacheq- ;
GMR	ag ^w myṛ	« terrain compris entre deux bras d'un cours d'eau, prairie » -tachelhit- ; agmir / aymir « limite entre deux terres, borne, bordure » -tarifit- ;
GM	gami	« frontière, limite » -tamacheq- ;
	game	« entre » -tamacheq-.

Sens 2 : « Puiser »

GM	agym / aḡym / ayyṃ	« puiser de l'eau, servir un plat » -tarifit- ; agym « puiser (dans un puits, avec un récipient ou un seau à la corde) » -tamacheq- ; ag ^w ym « puiser, vider » tachelhit, kabyle- ;
	aggam	« puisage » -tamacheq- ;
	ynagam	« puiseur » -tamacheq- ; anag ^w am « ouvrière qui puise de l'eau » -kabyle- ; aniyam « puiseur d'eau » -tarifit- ;
	asag ^w əm	« cruche en terre à puiser l'eau » -kabyle-.

Sens 3 : « Être plein »

GM	gəmməm	« regorger de, être garni (bassin) » -kabyle- ;
GMGM	ggəmgəm	« être plein à ras bord, grouiller, bouillonner » -kabyle- ;
	sgəmgəm	« bouillonner, produire un bruit sourd » -kabyle- ;
	agəmgum	« torrent, tourbillon, bouillonnement » -kabyle- ;
	gymgyṃ	« grommeler » -tarifit-.

La comparaison interdialectale, pour cette base, s'avère très intéressante pour deux raisons. En premier lieu, on remarque, dans toutes les langues, des évolutions phonétiques. Ces dernières peuvent compliquer en berbère l'identification de bases telles que nous les recherchons. En effet, il est difficile d'accepter d'intégrer à une base biconsonantique GM des termes ayant un sens proche mais contenant les radicales γM ou YM, comme nous l'avons fait pour aḡym et ayyṃ « puiser de l'eau ». Or, en prenant en compte les termes équivalents pour les autres variétés de la langue, il est facile de se rendre compte que le passage de [g] au [y] ou au [γ] est un phénomène régulier de cette variété, et de ce fait d'intégrer les termes contenant [γ] et [m] ou [y] et [m] à la base.

De la même manière, la comparaison interdialectale permet parfois de retracer l'histoire d'un sens et de pouvoir décider si un terme doit être intégré ou non à une base biconsonnantique. Par exemple, si l'on avait étudié exclusivement le tamacheq pour cette base GM, reliant des termes associés à l'idée d'étendue d'eau, il aurait été contestable d'intégrer le terme *ygdy*m « entourer de tous côtés ou d'un seul ». Cependant, en comparant avec un terme tarifit issu de la même racine : *agyddim* « rive d'un cours d'eau, talus, bord », on retrouve, dans le sens que véhicule ce dernier, la notion d'eau, qui a été évincée en tamacheq. La comparaison interdialectale permet donc de rétablir des liens sémantiques parfois relâchés ; ici, le lien sémantique entre l'étendue d'eau plane, la rive d'un cours d'eau et le fait d'entourer ne peut être mis au jour que par la comparaison interdialectale.

L'existence de bases biconsonnantiques reliées à une notion générale dans le lexique berbère est assez nette. Elles permettent de mettre en évidence des relations phonético-sémantiques qui ne peuvent être représentées par le concept de racine. Une étude approfondie de ces bases est tout à fait possible, les dictionnaires dont nous disposons étant de plus en plus nombreux et de bonne qualité.

Toutefois, il semble important de préciser qu'il manque encore, pour l'étude de ces bases biconsonnantiques, une théorie sémantique restrictive présentant des critères fiables pour autoriser ou exclure un terme d'une base biconsonnantique. Pour l'instant, si l'on se rend bien compte que des liens entre forme et sens ne peuvent être niés, les rapprochements sémantiques ont surtout pour fondement le bon sens : nous sommes obligés de ne prendre en compte que les sens qui peuvent être facilement rapprochés, ce qui donne forcément lieu à de nombreux oublis dans les bases ou à des rapprochements abusifs. Il faudrait, pour pallier ce manque, des informations étymologiques précises pour chacun des termes pris en compte, indications pour l'instant inexistantes. Ce manque nous limite donc dans la mise en place de ces bases, mais n'empêche pas d'en affirmer l'existence et de voir qu'elles concernent une grande partie du lexique.

3.3. *Bilan : les différentes bases biconsonnantiques dégagées*

J'avais mentionné, au début de cet article, différents articles isolés recensant quelques bases biconsonnantiques comparables à FR et GM. En mettant en commun ces différents travaux, l'existence des six bases biconsonnantiques suivantes, toutes hiérarchisées sémantiquement, semble établie pour le berbère :

- (b, r) ⇔ « Rondeur » (El Mountassir, 2004)
- (k, l) ⇔ « Déplacement, espace » (Aghali-Zakara, 2002 : 46)
- (l, γ) ⇔ « Mou, humide, tendre » (El Mountassir, 2002 : 194)
- (s, k) ⇔ « Pointu, Courbe » (Naït-Zerrad, 2000 : 88)
- (f, r) ⇔ « Séparation »
- (g, m) ⇔ « Étendue d'eau plane ».

Nous pouvons donc constater que ces bases de deux consonnes associées à un sens général, unités lexicales profondes qui permettent de mettre en évidence de nouvelles relations entre les lexèmes, couvrent une part importante du lexique. Ainsi, ce phénomène est moins anecdotique qu'on a pu le croire.

4. La théorie des matrices et des étymons, perspective inter-linguistique

Une ouverture vers une autre langue afro-asiatique a été très fructueuse.

En effet, il est intéressant de constater que ces bases biconsonnantiques que nous avons mises en évidence pour le berbère existent aussi en arabe, et qu'une théorie très complète a été

établie pour cette langue. Il s'agit de la Théorie des Matrices et des Étymons (TME) proposée par Bohas (1997, 2000) pour l'arabe.

Bohas montre qu'en utilisant le même raisonnement que celui employé pour définir la racine en arabe, on peut dépasser ce niveau, de la même manière que nous l'avons fait pour le berbère. Il propose alors une réelle restructuration du lexique, mettant en place trois niveaux : la matrice, l'étymon et le radical. La notion de racine est totalement exclue de cette structuration du lexique.

Dans la TME, le niveau le plus abstrait est celui de la matrice : il s'agit de combinaisons, non-ordonnées, de traits phonétiques auxquels est rattachée une signification primordiale commune, appelée « invariant notionnel » ; c'est-à-dire qu'il est possible, en arabe, d'identifier, pour de grands groupes de termes, des traits phonétiques communs reliés à un sens abstrait. Par exemple, un certain nombre de termes contenant les traits phonétiques {[+labial] [+coronal]} expriment quelque chose qui a à voir avec le fait de « porter un coup avec un objet tranchant ».

À un niveau un peu plus concret, sont présentés les étymons, composés binaires de phonèmes cette fois, consonantiques et non ordonnés, phonèmes comportant les traits phonétiques définis par la matrice, et qui sont liés à une charge sémantique développant l'invariant notionnel de la matrice, c'est-à-dire qu'ils précisent la notion abstraite proposée par la matrice. Par exemple, la matrice citée ci-dessus se trouve réalisée, en arabe, par l'étymon (b, t), ces deux phonèmes contenant respectivement les traits [+labial] et [+coronal], relié au sens général de « couper », spécifiant l'invariant notionnel de la matrice.

Enfin, au niveau le plus concret, on trouve le radical, élément qui apparaît quand on retire les affixes d'un lexème.

Cette théorie est assez avancée en arabe, et, ce qui est intéressant, pour nous, est que les étymons tels qu'ils sont définis par Bohas (1997 : 35) correspondent exactement aux bases biconsonantiques dont nous avons parlé pour le berbère. La TME est, au moins en partie, valide pour le berbère. Ainsi, il serait possible de parler d'étymons pour définir ces bases biconsonantiques berbères, et les études parsemées qui ont été proposées pour ces langues peuvent enfin être intégrées au sein d'une théorie éprouvée et trouvent, de fait, une légitimité supplémentaire. Nous n'avons pas le loisir, ici, de détailler les rapprochements entre les étymons berbères et arabes. On peut toutefois noter qu'en arabe, les deux phonèmes appartenant à un étymon peuvent être séparés par d'autres consonnes. Or, en berbère, si pour la base FR, les deux radicales communes étaient toujours accolées, on s'aperçoit, pour la base GM, qu'une consonne, voire deux, peuvent être intercalées entre les deux radicales phares. On observe cela, par exemple, dans le terme *gərməm* « former une nappe, stagner ». Ces unités biconsonantiques, mises en évidence de manière indépendante dans les traditions berbères et arabes, semblent vraiment équivalentes.

Le fait de bénéficier d'un cadre théorique permet en outre de faire des hypothèses sur l'ampleur de ce phénomène en berbère. En effet, si nous n'avons pas identifié de matrice pour le berbère, il est possible que l'on puisse en trouver lorsque les études sur les étymons seront plus avancées ; la TME propose aussi une analyse de l'homonymie intéressante, basée sur le croisement des étymons existants : il faudra vérifier si une telle analyse est possible pour le berbère.

Toutefois, si cette théorie qui fonctionne pour l'arabe est applicable au berbère dans son ensemble, j'émettrais une réserve quant à l'abandon total de la notion de racine. L'utilisation de cette notion est, selon moi, tout à fait justifiée en berbère, ne serait-ce que parce qu'elle se base sur un raisonnement identique à celui qui nous a servi à identifier les étymons. Cependant, tout comme l'étymon coexiste avec le radical et la matrice dans la TME, il peut

très bien cohabiter de la même manière avec la notion de racine, l'étymon se situant simplement à un niveau plus abstrait.

5. Conclusion

Ainsi, au sortir de cette étude, nous avons confirmation que la proposition de différents auteurs à propos de la possibilité d'un autre type d'organisation du lexique berbère était justifiée : il existe bien des unités lexicales plus fines que la racine, établies grâce au même raisonnement, reliant, dans des ensembles dérivatifs, deux phonèmes communs à un sens général constant. Les notions de racine et de schème, si elles permettent de décrire le lexique, sont impuissantes pour mettre en lumière ces relations phonético-sémantiques profondes du lexique. Cette étude nous a aussi permis de constater, en mettant en commun les différents travaux déjà effectués et en les complétant, que l'ensemble du lexique semble être régi dans son ensemble par ces unités biconsonantiques profondes ; ces « étymons » auraient donc un rôle central dans l'organisation du lexique berbère, et, s'il est encore difficile de les identifier avec certitude, ils ne doivent pas pour autant être négligés. Le fait de pouvoir reconnaître un fonctionnement équivalent dans une langue sœur du berbère conforte encore l'existence de ces bases et permet d'adopter un cadre théorique qui manquait à cette étude.

Références

- AGHALI-ZAKARA, Mohamed (2002). Racines homophones en berbère : cas de KL, in *Articles de linguistique berbère, Méorial W. Vycichl*, Naït-Zerrad, Kamal (Ed). Paris : L'Harmattan 43-56.
- ALOJALY, Ghoubeïd, MOHAMED, Ghabdouane & PRASSE, Karl-G (2003). *Dictionnaire touareg-français*. Copenhagen : Museum Tusculanum Press.
- BOHAS, Georges (1997). *Matrices, étymons, racines, éléments d'une théorie lexicologique du vocabulaire arabe*. Paris-Louvain : Peeters.
- BOHAS, Georges (2000). *Matrices et étymons, développements de la théorie*. Lausanne : Editions du zèbre.
- CADI, Kaddour (1987). *Système verbal rifain, forme et sens*. Paris : SELAF.
- COHEN, David (1988). Les langues Chamito-sémitiques, in *Les langues dans le monde ancien et moderne*, Perrot, Jean (Ed). Paris : CNRS, 9-30.
- COHEN, David (1993). Racines, in *À la croisée des études libyco-berbères*, Drouin, Jeannine & Roth, Arleth (Eds). Paris : Geuthner, 161-175.
- DALLET, Jean-Marie (1982). *Dictionnaire kabyle-français*. Paris : SELAF.
- DESTAING, Émile (1920). *Lexique français-berbère, Étude sur la tachelhit du Sous*. Paris : E. Leroux.
- EL MOUNTASSIR, Abdallah (2002). La notion de "mou, humide, tendre" en berbère : l'étude de la base Lγ, in *Articles de linguistique berbère, Méorial W. Vycichl*, Naït-Zerrad, Kamal (Ed). Paris : L'Harmattan, 181-196.
- EL MOUNTASSIR, Abdallah (2004). Étude d'un champ morphosémantique en berbère : le cas de l'étymon BR, in *Actes du colloque Bayreuth-Frankfurt-Leidener Kolloquium zur berberologie*.
- GALAND, Lionel (1988). Le berbère, in *Les langues dans le monde ancien et moderne*, Perrot, Jean (Ed). Paris : CNRS, 207-242.
- NAÏT-ZERRAD, Kamal (2000). Autour de la base morphosémantique SK en berbère, in *Études berbères et chamito-sémitique, Mélanges offerts à Karl-G. Prasse*, Chaker, Salem & Zaborski, Andrzej (Eds). Paris-Louvain : Peeters, 295-302.
- SERHOUAL, Mohamed (2002). *Dictionnaire Tarifit-Français, Thèse de doctorat dirigée par F. Bentolila et M. Taïfi*. Tétouan : Université Abdelmalek Esaâdi.
- TAÏFI, Miloud (1992). *Dictionnaire tamazight-français*. Paris : L'Harmattan.

POUR UNE ÉTUDE DE L'APOSTROPHE INJURIEUSE

Iuliana-Anca MATEIU

iuliamateiu@yahoo.com

Faculté des Lettres, 31 rue Horea

400320 Cluj-Napoca

Roumanie

Abstract : Through a comparison of the insulting vocative to other possible examples of vocatives (titles, tender words or denotative nouns in the vocative) the functioning of insulting nouns will be examined, as well as the semantic restrictions concerning the vocative construction itself.

1. Introduction

Malgré la relativisation de la définition des injures par les pragmaticiens, qui mettent l'accent sur leur effet perlocutoire, sur leur réception qui les confirmerait comme telles ou les infirmerait, nous considérons que leur formulation et le matériau langagier (lexical, syntaxique) utilisé dans ce sens n'en sont pas moins importants. En fait, l'effet d'injure est le résultat de l'interaction de plusieurs composantes, dont certaines proprement linguistiques, d'autres extra- ou paralinguistiques.

Dans ce qui suit, nous allons analyser le rôle de chacune de ces composantes et surtout une des constructions syntactico-énonciatives qui sert l'injure, à savoir l'apostrophe, considérée dans deux de ses réalisations : comme apostrophe neutre vs injurieuse.

A travers une comparaison d'apostrophes construites sur un N titre, sur un N insultant ou hypocoristique ainsi que sur des N dénotatifs modifiés ou non, nous étudierons tant le fonctionnement d'un type particulier de N (les N insultants) que les contraintes sémantiques qui pèsent sur la construction même de l'apostrophe.

2. La composante lexicale

La composante lexicale ne suffit pas – comme on pourrait le conclure du fait que le terme *injure* désigne couramment toute une série de mots offensants, recueillis dans des dictionnaires d'injures – à définir une injure. Aussi parle-t-on plutôt d'une classe ouverte, mal définie de N insultants, qui comporte un fonds stable de N – les « Noms de qualité » (NQ) (Milner, 1978) – appelés ainsi d'après leur valeur axiologique (le plus souvent négative), et des créations plus ou moins inédites, de nature métaphorique ou métonymique, reconnues comme insultantes grâce à leur inadéquation référentielle autant qu'à l'intonation et à des indices paralinguistiques (mimogestualité) et extralinguistiques (situation conflictuelle). Inadéquation référentielle, mais pas situationnelle, car généralement cet écart de langage et de conduite n'est qu'une sanction par laquelle l'injuteur essaye de corriger une attitude de l'injurié qui lui crée du désagrément.

Issues d'un vif mouvement affectif, les injures sont de véritables mises en scène, où la parole n'est qu'un des moyens d'expression. Ainsi peut-on expliquer le fonctionnement comme injures de termes sans aucun contenu appréciatif, des N ordinaires, qui sous le coup d'une énonciation particulière, acquièrent une signification nouvelle : celles de marques de rejet.

C'est là une conséquence de la recherche d'expressivité caractéristique de tout discours affectif. La contrepartie des injures, les termes caressifs (hypocoristiques/mots tendres) sont tout aussi nombreux, car la relation intime qu'ils sont censés traduire et renforcer dispense en

partie des contraintes sociales et permet ainsi une plus grande liberté d'expression envers l'autre.

3. La composante prosodique

Les auteurs qui ont signalé la relativité dans la définition des appellatifs affectifs ont également remarqué comme un cas extrême la possibilité d'employer un même terme avec une valeur ou une autre, et ce, même si l'un des emplois est plus fréquent que l'autre. Ce qui fait la différence, et porte la signification, c'est l'intonation.

Dans certains cas (avec des NQ : *salaud/ordure/andouille*, etc.), l'intonation ne fait que souligner le sens du lexème (devenant en fait redondante), dans d'autres, elle peut le contredire ou seulement l'orienter, avec la participation d'autres éléments, vers une appréciation +/- . Une injure lexicale dite avec une intonation tendre peut être interprétée comme un hypocoristique ou paraître au moins ambiguë. Un tel emploi relève de l'ironie, que Berrendonner (1981 : 222) explique en termes de « paradoxe argumentatif ». Favorisé par « le caractère *pluri-codique* de la communication », le paradoxe argumentatif consiste en la convocation, dans un seul acte d'énonciation, de deux arguments contradictoires : ce que dit l'énoncé de son énonciation (par ses variables argumentatives) se trouve contredit par ce que l'énonciation dit d'elle-même (par des « symptômes » mimogestuels).

S'adresser à un ami avec la formule :

1. – Salut, *vieux débris, cuistre vain !* (*apud* Berrendonner),

mais un charmant sourire aux lèvres, ne sera plus interprété comme une injure, tout au plus comme une énonciation ambiguë, en fonction des susceptibilités de l'interlocuteur.

4. La composante syntactico-énonciative

La construction syntactico-énonciative où s'insèrent ces lexèmes est à son tour significative des rapports entre les partenaires discursifs, et parfois décisive pour l'interprétation du nom. Certaines tournures se sont spécialisées dans l'expression de l'injure, de sorte qu'elles convertissent en N insultants des N non axiologiques, voire des axiologiques (+) (cf. *espèce de morphème !*; *Pd N1 de N2 : ce génie d'Einstein « ce fou d'Einstein »*). Parmi celles-ci, l'apostrophe est à l'honneur dans la formulation des injures.

Étant donné d'un côté la visée perlocutoire des injures – de rejeter le référent (*tu/il*, voire *je*) en le rabaissant – de l'autre, le caractère direct de la désignation dans le cas de l'apostrophe, on peut comprendre pourquoi les injures préfèrent cette formule et pourquoi la formule elle-même peut en générer de nouvelles.

4.1. *Apostrophe injurieuse vs apostrophe neutre*

Les NQ partagent l'emploi en apostrophe avec quelques N fonctionnant comme *termes d'adresse* (TA) et avec les N propres (Npr). Des N sémantiquement analogues à ceux qui font office de titres en sont cependant exclus (cf. (venez), *docteur/général !* vs *(venez), *médecin/étudiant !*) et même certains NQ : les qualifiants [-animé] (cf. *merveille, horreur*) et certains [+animé] tels : *(venez), *amour/diable !*.

Cette restriction formulée par Milner au sujet des N pouvant figurer en apostrophe est remise en cause d'abord par Ruwet (1982), ensuite par Lagorgette (1998), qui lui reprochent d'ériger en norme un jugement personnel, puisqu'il ne travaille pas à partir d'un corpus d'exemples, mais d'après sa propre compétence de locuteur natif. Cela mène inévitablement à des divergences dans les jugements de (a)grammaticalité à propos d'un emploi ou d'un autre.

Pour ce qui est des N [-animé], [-humain], leur exclusion de l'apostrophe ne se justifie que si l'on ne tient pas compte de la métaphorisation qui joue un rôle important dans la

dénomination. Les meilleurs contre-exemples pourraient être justement les apostrophes injurieuses, qui sont constituées, pour une large part, de N [-animé, -humain], ou seulement [-humain] employés de manière figurée (métaphorique ou métonymique) :

2. Coupable moi ? Vous osez m'accuser ? *Misérables vers de terre ! Cornichons ! Cloportes ! Négriers ! Coléoptères ! Sapajous ! Topinambours ! Vermicelles ! [...]* (Hergé)

Il s'agirait dans ce cas d'*apostrophes personnifiées*¹, de même que dans des exemples tels que :

3. *Ô lac ! rochers muets ! grottes ! forêt obscure ! [...]*
gardez, *belle nature*,/ Au moins le souvenir ! (Lamartine),

lesquels n'ont pas nécessairement une valeur qualitative (apostrophes neutres).

Au-delà de leur fonction ludique et poétique, ces bordées d'injures témoignent de la spécificité d'un type de discours qui, grâce à sa forte dramatisation, peut user un peu plus librement des ressources lexicales.

4.2. *Les injures, termes d'adresse [+axiologiques négatifs]*

Si l'on insiste tant sur le potentiel locutoire du référent désigné par le N en apostrophe, c'est à cause de la fonction exclusivement interpellative de l'apostrophe.

D'autre part, tandis qu'une partie des injures (métaphoriques/métonymiques) et d'autres apostrophes acquièrent le trait [+locutoire] occasionnellement, à la faveur d'un emploi figuré (comme personnifications), une autre partie, de même que les titres et les Noms propres le présentent comme un trait stable et en déduisent même leur principale fonction discursive : celle de termes d'adresse. Le fait que, du point de vue morphosyntaxique, les injures se comportent comme les autres termes d'adresse, expliquerait pourquoi la plupart des auteurs négligent leur principale différence : si leur suppression modifie le sens et la fonction de la phrase, celle d'un terme d'adresse ne l'affecte pas de manière significative, comme il ressort des exemples suivants :

4. Mais non, j'ai pas honte, *idiot*, tu me mets en valeur, au contraire. La belle et la bête. (Sarraute, *Dis, tu m'aimes ?*, p. 114) : contestation + injure
5. Mais non, j'ai pas honte, *Marie/maman/ma chère*, tu me mets en valeur, au contraire : contestation
6. Mais non, j'ai pas honte, tu me mets en valeur, au contraire : contestation

L'appellatif détonne et impose son sens au contexte, surtout quand il est négatif. Dans un contexte plus large, la suppression d'une injure peut passer inobservée, comme celle d'un appellatif neutre, car le sens repose sur plusieurs éléments (éventuellement d'autres injures) :

7. L'armée française, elle te pisse au cul, (*eh, planqué*) ! répond le galonné. (Cavanna, *Les Russkoffs*, p. 66).

La suppression des termes d'adresse peut devenir elle-même signifiante, à un niveau supérieur : celui de l'échange. Étant donné qu'ils définissent les statuts et les rôles réciproques du locuteur et de l'allocutaire, leur usage inadéquat bouleverse la relation avec l'interlocuteur, qui se voit redéfini autrement qu'il ne pensait être et, par conséquent, peut se sentir injurié. Quant à leur absence, elle peut signifier le refus du locuteur de définir son rapport à l'autre, ce qui pourrait également passer pour une injure. En fait, comme le remarque Kerbrat-Orecchioni (2002 : 32) :

En France, il semble bien en outre [...] que l'on assiste aujourd'hui à une raréfaction notable de l'emploi des noms d'adresse : dans bien des situations communicatives, la salutation et le remerciement ne s'accompagnent plus

¹ LAGORGETTE, Dominique (2001).

automatiquement d'un nom d'adresse comme le recommandent grammaires et traités de savoir-vivre (en revanche, le nom d'adresse apparaît volontiers aux côtés d'un reproche, d'une protestation ou d'une réclamation, c'est-à-dire qu'il a souvent une connotation polémique).

Autrement dit, les variations diachroniques en matière d'adresse vont dans le sens d'une réduction des occurrences, voire d'un réinvestissement des appellatifs comme marques de conflit, donc plus si éloignés des injures.

On signalait *supra* l'emploi de plusieurs termes dans le sens de *termes d'adresse* : *apostrophe*, *vocatif*, *appellatif*. Ce sont là des dérivations métonymiques à partir de la fonction qu'assument ces termes. Afin de pouvoir identifier les principales caractéristiques de leur construction, nous allons revoir d'abord les motivations et aspects évoqués par d'autres analystes qui ont choisi de parler du *vocatif* (1), ou bien de l'*apostrophe* (2) ou de l'*appellatif* (3).

Le terme *vocatif* nomme le cas désigné à exprimer l'adresse à un allocataire humain, mais pas exclusivement (animaux, objets, concepts) et marqué par des terminaisons spécifiques dans les langues à déclinaison, comme le latin ou le roumain.

Le Goffic (1993) et Frontier (1997) distinguent deux sortes de vocatifs : injurieux vs proprement dit. Si le premier sert à caractériser (qualifier) l'allocataire, avec une forte dose de subjectivité, le second sert à l'identification de l'allocataire et peut éventuellement s'accompagner de la mention de qualités annexes permanentes, non liées à la prédication de l'énoncé (cf. *cher Paul*) ou même se faire par une simple qualité marquée par un adjectif (cf. *Petit*, viens ici !). Lorsque associé à une phrase, le vocatif proprement dit a une place *a priori* pertinente (plus il est situé en début d'énoncé, plus il a de chances de jouer un rôle dans l'identification de l'allocataire), tandis que le vocatif injurieux a une place variable. Lorsque les constructions à NQ manifestent leur entière autonomie énonciative, elles représentent des « phrases nominales attributives » (Le Goffic, 1993) à un ou à deux termes qui apostrophent et caractérisent l'allocataire.

Lefeuve (1999) distingue l'apostrophe de la phrase averbale à sujet implicite dont les illustrations nous ramènent aux injures interpellatives (cf. *Monstre !/Petit coquin !*).

Alors que ces phrases averbales impliquent une prédication, qu'elles forment des jugements, l'apostrophe n'est qu'un « membre de phrase », classé parmi les « modulations » qui accompagnent souvent et permettent de reconnaître une phrase injonctive. L'apostrophe peut servir à identifier l'allocataire pour que le locuteur puisse agir sur lui (cf. *Garçon*, deux demi-tasses !), soit à maintenir le contact (fonction phatique), une fois que l'identification de l'allocataire est établie par la situation. Enfin, lorsque le locuteur place son discours sous le patronage d'une instance, l'apostrophe n'est qu'une prise à témoin (cf. *Mon Dieu ! de l'or à tout prix ! se disait L.*).

Cependant, la distinction entre apostrophe et phrase averbale à sujet implicite n'est plus si évidente quand un NQ est suivi, sans forte séparation, d'une phrase verbale où la 2^e personne apparaît sous la forme d'un verbe à l'impératif ou d'un pronom personnel (cf. Stendhal (1830), *Le Rouge et le noir* (chap. IV) : *Ah, monstre ! s'écria Julien à demi-haut, et des larmes généreuses vinrent mouiller ses yeux... Ah, petit gueux ! pensa-t-il, je te revaudrai ce propos.*). L'existence de cas pareils amène l'auteur à parler d'un continuum entre l'apostrophe et respectivement la phrase averbale à un terme, qu'elle représente d'ailleurs sur un axe du genre :

Monsieur	<i>petit gueux</i>	<i>monstre !</i>
Apostrophe		phrase averbale à sujet implicite

Les termes qui constituent les prédicats de ces phrases averbales sont pour la plupart des SN axiologiques. Des termes objectifs peuvent constituer la base prédictive de phrases averbales existentielles, voire des averbales attributives, mais leur emploi apparaît plus contraint que les précédents.

Selon Curat (1999 : 202), les *invectives* ressemblent aux apostrophes, mais sans s'y confondre, car si elles représentent des « étiquettes lancées à l'auditeur, qui le classent en portant sur lui un jugement de valeur », les apostrophes classent en effet l'auditeur, mais leur but n'est pas cet étiquetage en soi.

Selon Lagorgette, le terme d'*appellatif* renvoie au départ à tout terme de dénomination pour ce qui n'est pas humain ; ensuite il connaît une extension de sens, pouvant renvoyer à des N communs comme à des N propres, présents dans le discours direct ou dans le récit.

En fait, la spécificité de l'apostrophe injurieuse se révèle mieux lors d'une hiérarchisation des N selon leur capacité à assumer la fonction vocative. Selon Clarinval², le seul – à notre connaissance – à suggérer une telle hiérarchie, après un N commun [+titre] tel *Sire* destiné exclusivement à un emploi vocatif, viennent les N propres et d'autres N communs [+titre], lesquels admettent aussi un emploi désignatif (cf. *Dupont/ Monsieur*, vous oubliez votre chapeau ! / Je ne me rappelle plus qui est *Dupont/ ce monsieur*. ; *C'est monsieur Dupont*). Les N communs qui ne présentent d'autres traits que [+humain] seraient les derniers, car tous ne sauraient pas s'employer au vocatif (cf. *Homme*, viens !).

Nous remarquons cependant que des modifications telles que la qualification par un adjectif épithète ou par un complément prépositionnel rendent leur emploi au vocatif plus naturel. Ces emplois de même que l'emploi non contraint de N similaires en français (cf. *Femme*, monceau d'entrailles, pitié douce...³), voire de ce N même (et d'autres) en roumain, langue parente (cf. *Vino*, *omule* !), nous font penser qu'il s'agit moins d'une explication sémantico-pragmatique, qui devrait s'appliquer dans d'autres langues aussi, que d'une question d'usage.

Considérons, par exemple, les constructions suivantes :

- a. Nom commun [+humain] :
 - 8. *Omule*, învață să fii bun ! → ? *Homme*, apprends la générosité !
 - b. Nom commun [+humain] + génitif :
 - 9. Dă-mi odată pace, *omu'/femeia naibii* ! → ? Arrête de m'embêter, *homme/ femme du diable* !
 - 10. Ce te-mpingi așa, *omu' lu' Dumnezeu* ! → ? Qu'est-ce qui te prend à pousser comme ça, *homme de Dieu* ?
 - c. Nom commun [+humain] + de + SN :
 - 11. *Om de rea credință* ! → (...) *homme de peu de foi* ! (San-Antonio)
 - 12. *Om de doi bani/ de nimic/ de rahat (ce ești)* ! → ? *Homme de quat' sous !/de rien /de merde ; Femme de peu* ! (Céline apud Gordienne)

Si dans (A), le nom *om / homme* est un simple terme d'adresse qui, d'un côté, sert à attirer ou bien à maintenir l'attention de l'allocutaire sur une communication que le locuteur va / est en train de lui faire; et de l'autre, applique au référent une dénomination (valable de manière) unique et permanente, dans (B), où le N commun est complété d'un génitif, l'interpellation s'accompagne de l'expression d'un sentiment – négatif (d'impatience ou de contrariété), qui rappelle le juron, puisque le complément est une évocation blasphématoire (*de Dieu* ou *du diable*). L'intonation participe de cette expression affective. La qualification semble l'emporter ici sur la dénomination et elle s'appuie obligatoirement sur le cotexte (le contenu de l'énoncé). Les dernières constructions (C) servent, comme les autres, à interpeller, mais

² Apud Lagorgette (1998).

³ Cavanna, *Bête et méchant*, p. 76.

surtout à qualifier l'allocutaire (ou, éventuellement, le délocuté, cf. *infra*). Le N reçoit cette fois un complément de prix – qui exprime la valeur du référent. Celle-ci étant des plus basses, les SN se chargent d'une valeur axiologique négative et finissent par fonctionner comme injures. À la différence des appellatifs précédents, celles-ci peuvent apparaître toutes seules, étant équivalentes à une prédication. Parmi ces compléments de prix, on trouve en français : *de balle/ de merde/ de première bourre ; de malheur ; de mes pieds/ de mes deux/ de mes couilles/ de mon cul/ de mes fesses*. Construits sur des gros mots (scatologiques ou sexuels), la qualification qu'ils apportent ne peut être que dégradante et les GN où ils apparaissent, des injures. En fait, à part ces cas où le premier N est un N neutre que le complément dévalorise, il y en a où N1 est un N insultant que le complément renforce, comme dans les exemples ci-dessous :

13. *Garçon de ma merdre*, si je t'en croyais je ferais rebrousser chemin à toute l'armée. (Jarry *apud* Gordienne)

14. - C'est fini, oui ou merde ? Eh, nous on bosse, nous demain ! *Salopes de merde ! Poufiasses ! Sauvages ! [...]* (Cavanna, *Les Russkoffs*, p. 259)

En français, on devrait ajouter également des constructions à complément de lieu (origine) (cf. *de barrière/ de banlieue/ de 2^{ème} zone/ de chiotte/ de sous-préfecture/ de ruisseau/ de poubelle/ de pissotière*), qui attribuent au référent de N1 une origine toujours dégradante, aboutissant au même effet que les précédents (génération ou renforcement d'une injure) :

15. Va te faire baguer le nœud, eh, *baba de pissotière !* (Perret).

L'observation du comportement de ces diverses catégories de N en vocatif nous conduit aux hypothèses explicatives suivantes :

Si les *N titres* (*Docteur, professeur, mon général* ou bien *Monsieur, Madame*) ou les termes relationnels (*maman, frangine, fiston*) peuvent apparaître en vocatif, c'est parce qu'ils sont par excellence des *mots d'interaction* : ils servent à instituer un rapport, social ou affectif, entre les interlocuteurs, en désignant le rôle que le locuteur reconnaît et impose à l'autre, et indirectement son propre rôle. Ils expriment donc une catégorisation stable, reconnue par une micro-communauté (familiale, professionnelle).

Au pôle opposé se placent les *injures*, catégorisations accidentelles, dictées par le contexte situationnel, éventuellement décrit dans le contenu de l'énoncé. Elles représentent en quelque sorte une *parodie des termes d'adresse admis*, puisqu'elles redéfinissent le référent à chaque fois. D'une part, elles communiquent à l'allocutaire un jugement, toujours renouvelé, donc à la fin injustifié, du locuteur à propos de lui. D'autre part, en annulant la distance sociale, elles représentent une forme de contestation des usages sociaux. Cependant, elles intéressent l'interlocuteur autant que les appellatifs conformes ; par contraste, on dirait, mais aussi par leur part de subjectivité, qui leur assure un supplément d'informativité.

Entre les deux, se situent les N dénotatifs ordinaires, qui renvoient à une identité stable, reconnue, une nature. Aussi ne sauraient-ils s'employer en apostrophe seuls, comme les autres. N'ayant ni une signification sociale ni une signification affective comme les termes d'adresse (titres, termes relationnels, N propres, hypocoristiques ou injures), ils ont besoin d'un contexte linguistique qui accentue justement leur sème [+humain, +locuteur], ne fût-ce que pour le contester (cf. les exemples sous 2° et 3°). Lorsque son interlocuteur semble déchoir de sa qualité d'homme, par une conduite inappropriée, ou bien lorsqu'il semble ne plus comprendre le langage humain, ses appels répétés, le locuteur tente de le remettre à sa place. Ou bien, en appelant une femme par son nom, on inscrit automatiquement celui-ci dans un nouveau réseau de significations : la féminité est valorisée ou dévalorisée (en fonction du contexte) et son évocation signifie à l'autre qu'elle doit l'assumer comme le rôle auquel renvoie le titre ou le terme relationnel. Si le locuteur choisit, de tous les appellatifs

possibles, celui de *femme*, dans un contexte tel que *Femme*, j'amène des invités / *Nevastă*, ne vin musafiri., c'est bien pour l'inviter à jouer son rôle de maîtresse de maison, de femme⁴.

Plus un appellatif est général, plus l'énoncé doit contenir d'informations qui restreignent son extension et, en fin de compte, justifient son emploi. Deux types de restrictions sont possibles : par la prédication qui suit, en général, ces N en apostrophe (cf. 1^o et 2^o), ou bien par des qualificatifs (adjectifs : *homme cruel* ; compléments de prix : *femme de peu/ homme de malheur*; *om de doi bani/ de rahat* ; génitif à valeur blasphématoire : *omu' lu' Dumnezeu*), grâce auxquels le SN en apostrophe devient lui-même une prédication fonctionnant au besoin pour son propre compte, avec le seul appui du contexte extralinguistique.

On remarquera également que l'ajout d'un prédéterminant (l'article défini, mais surtout l'adjectif possessif : *mon/ma/mes*) peut avoir le même effet qu'un qualificatif : le possessif, rattachant le référent au monde du locuteur, se charge d'exprimer l'affectivité, avec l'intonation. Des interpellations comme *ma fille*, *mon garçon*, *mon fils*, *mon frère*, *ma soeur* sont plus naturelles que *fille*, *garçon*, *fils*, *frère*, *soeur*, à cause justement de leur nuance affective. Cela est encore plus évident dans le cas de leurs correspondants familiaux, dont certains portent des suffixes diminutifs ou augmentatifs, à valeur affective (cf. *fifille*, *frérot*).

En roumain, on observe le même phénomène chez des N peu pertinents comme interpellatifs : *studentule* (l'étudiant), *artistule* (l'artiste...) ou le féminin *artista*, etc., à moins qu'ils ne prennent une nuance ironique, comme dans les contextes :

16. Hai, *artistule*, la tablă ! «Viens faire ton numéro, l'artiste ! » (à un élève qui fait le pitre) (défi, mépris)

ou bien d'une valeur affective positive, comme c'est le cas dans des énoncés du genre :

17. (Vino), *studentu' / artistu' lu' mama* ! «Viens, mon artiste !/ l'artiste !» (attendrissement, admiration, fierté)

Le rôle de l'adjectif possessif est assumé ici par le génitif (*lui mama/tata*) et la sphère des possibilités est restreinte à ces deux termes relationnels. Les parents, et surtout la mère, sont les plus susceptibles de s'approprier ainsi, amoureusement, les qualités et défauts (mineurs) de leur descendance (cf. *deșteptu' / frumosu' / prostănacul / prostuțu' lu' mama* ! « mon petit malin / mon beau garçon / mon petit con ! », qui rappellent facilement d'autres expressions affectives telles que *dragul / scumpu' / odorul / puțul mamei* « mon chéri / mon petit ! », mais * *tâmpitu' / ticălosu' lu' mama / mamei* ! « mon idiot / mon salaud ! »).

À l'encontre des titres et autres termes d'adresse neutres, qui ont avant tout une fonction dénomminative, les dernières formules servent surtout à caractériser, la fonction conative passant au second plan. L'ambivalence de ces constructions, qui peuvent servir en égale mesure d'interpellatifs, ou de désignatifs à la 3^e personne, n'y est peut-être pas pour rien.

En l'absence du génitif, qui leur confère une valeur affective-appréciative positive, ces mêmes N prennent une valeur contraire : d'appellatifs ironiques, voire injurieux. Dans

18. *Deșteptule* ! (Ai văzut ce-ai făcut ?) → Tu vois maintenant ce que tu as fait, l'astuce ? / *Gros malin* ! (T'as vu ce que t'as fait ?)

19. *Deșteptu'* ! (Ai văzut ce-a făcut ?) → *Le gros malin* ! T'as vu ce qu'il a fait ?

le nom voit s'inverser sa valeur axiologique (du positif au négatif), ce qui lui permet d'exprimer un reproche, une critique. Par antiphrase, le compliment tourne à l'injure (*deșteptu(le) / (le) gros malin* ! signifie plutôt *tâmpitu(le) / (l') idiot, (le) connard !, Ducon*), comme le prouvent ces réponses à (18) : *Deștept ești tu cu (mu)mă-ta* ! (l'équivalent du français *X toi-même !*), ou bien le reproche : *De ce mă-njuri ?* « *Qu'est-ce qui te prend à m'insulter ?* ».

⁴ Curat (1999).

En roumain, l'ironie devient plus mordante si l'on renforce le nom par redoublement : *Deșteptu' deșteptilor !* « *Le malin des malins !* ». L'effet superlatif peut être obtenu aussi par la mention du milieu où le référent est unique par sa « qualité » : *deșteptu' lumii/pământului !*, quand l'individu n'est désigné par le nom même de la qualité qu'il possède et qu'il finit par incarner : *deșteptăciunea pământului !* « *sagesse de la terre !* ».

La proportion des qualifications négatives, respectivement positives dans cette construction (9 à 1), dans les exemples cités par Lüder (1996) (cf. *faurul lumii* « *le forgeron du monde* » vs *Bată-te varga lui Dumnezeu, hâzenie a pământului !* « *Que Dieu te punisse, horreur de la Terre !* ») serait-elle le signe d'une spécialisation de la construction même dans l'expression au superlatif d'un défaut ? Cela ne diminuerait pourtant pas l'importance de l'appellatif. Le fait que le terme ait toujours une valeur dépréciative, ironique, lors d'un emploi désignatif (cf. L-am întâlnit pe *deșteptul ăla* și la bibliotecă. « J'ai rencontré *ce malin-là* à la bibliothèque aussi. »; *Care deștept a făcut asta ?* « *Quel malin a fait ça ?* »), prouve qu'un qualifiant, même appréciatif, ne saurait servir de dénomination sans changer de valeur. Le contenu subjectif appréciatif de ces N réduit leur potentiel dénotatif. Aussi ne sauraient-ils fonctionner comme désignations en dehors d'un consensus comme celui auquel on aboutit lors de l'invention d'un surnom, par exemple. L'emploi de tout nouvel appellatif dont la dénotation n'est pas stabilisée par un tel consensus et qui ne sert pas à exprimer une appréciation affective sera interprétée comme une tentative de redénomination, une parodie de baptême et, selon le contenu, comme une véritable injure ou une simple ironie.

En combinant l'identification du référent avec sa convocation directe dans l'énoncé, l'apostrophe sert parfaitement les injures, qui recatégorisent le référent sur un échelon plus bas et qui ont d'autant plus de force qu'elles sont directes. Cette idée de recatégorisation de l'allocutaire explique à rebours la pertinence des N insultants en apostrophe.

Références

- BERRENDONNER, Alain (1981). *Éléments de pragmatique linguistique*. Paris : Ed. de Minuit.
- CURAT, Hervé (1999). *Les Déterminants dans la référence nominale et les conditions de leur absence*. Genève-Paris : Librairie Droz.
- FRONTIER, Alain (1997). *La Grammaire du français*. Paris : Belin.
- GORDIENNE, Robert (2002). *Dictionnaire des mots qu'on dit « gros », de l'insulte et du dénigrement*. Paris : Ed. Hors Commerce.
- KERBRAT-ORECCHIONI, Catherine (2002). Adresse, in *Dictionnaire d'analyse du discours*, Charaudeau, Patrick & Maingueneau, Dominique (coord). Paris : Ed. du Seuil, 32.
- LAGORGETTE, Dominique (1998). *Désignatifs et termes d'adresse dans quelques textes en moyen français*, thèse Parix X Nanterre.
- LAGORGETTE, Dominique (2001). *Les axiologiques négatifs sont-ils une classe lexicale ?*, communication au colloque *Représentation du sens linguistique*, Université de Bucarest.
- LEFEUVRE, Florence (1999). *La Phrase averbale en français*. Paris : L'Harmattan.
- LE GOFFIC, Pierre (1993). *Grammaire de la phrase française*. Paris : Hachette.
- LÜDER, Elsa (1996). *Procedee de gradație lingvistică*. Iași : Editura Universității A. I. Cuza.
- MILNER, Jean-Claude (1978). *De la syntaxe à l'interprétation. Quantités, insultes, exclamations*. Paris : Ed. du Seuil.
- RUWET, Nicolas (1982). *Grammaire des insultes et autres études*. Paris : Ed. du Seuil.

LA DIMENSION INDIVIDUELLE DE LA PAROLE : APPROCHES LINGUISTIQUE ET COGNITIVE

Séverine MORANGE

Severine75@free.fr

Recherche sur le français contemporain (EA 1483)

Université de Paris 3 ; 13, rue de Santeuil, 75005, Paris

Abstract : The main objective of this study was to scan and codify the intra and inter-individual variations in the spoken French of three Parisian senior citizens. Interviews with the three speakers were analysed and recordings of their voices were submitted to ordinary listeners, in order to identify the dimensions of speech that are relevant in characterising individuals. Prosodic variation and voice categorization were examined.

1. Introduction

Sous son versant oral, la parole apparaît comme une manifestation pluriparamétrique par excellence. Elle est sans doute le mode d'échange privilégié au travers duquel les formes de la subjectivité du sujet se révèlent, se font entendre ou s'éprouvent. Ceci tant au niveau de la production que de la perception, et tant du point de vue du linguiste que du sujet ordinaire. La présente contribution trouve sa source dans notre recherche de doctorat (Morange, 2005a), dont la variation linguistique individuelle repérée à partir de l'analyse de la parole, constitue le thème central.

1.1. Problématique

Notre intérêt pour l'étude de la parole tient précisément à ce qu'elle dévoile de l'individu, à travers les mécanismes (« figures imposées ») de la langue en tant que donnée partagée. Nous considérons que la parole proférée est intrinsèquement et consubstantiellement incarnée, individuée et individualisée, par-delà la norme sociale intériorisée de la langue employée.

Le premier axe de cette bipolarité de la parole qu'est l'individualité, en contraste avec l'usage d'une norme, représente donc le point focal de notre étude. Il constitue également un élément original, selon nous, au regard d'enjeux contemporains touchant spécifiquement à la question de la langue et de l'industrie vocale, davantage concernés par l'universalité que par l'individualité.

1.2. Hypothèses

Le projet a pour origine deux questions essentielles :

- qu'est-ce qui fait l'individualité d'une parole ?
- comment appréhender « le grain », l'effet produit par l'écoute d'une voix ?

Envisageant une recherche de type monographique, nous postulons que le repérage de niveaux d'invariants et de variations individuelles de sujets parlants donnés peut s'opérer dans la référence à différents domaines disciplinaires, l'un situé du côté de la production (l'analyse prosodique ici), l'autre du côté de la perception (l'analyse catégorielle ici).

2. Domaines de recherche

2.1. Diversité de cadrages théoriques

Le champ prosodique est abordé :

- d'un point de vue discursif (Morel & Danon-Boileau, 1998 ; Morange, 2003), et à travers le concept de « mélisme » (Caelen-Haumont & Bel, 2003 ; Caelen-Haumont, 2004 ; Morange 2004b), qui permet de mettre en évidence l'importance de propriétés mélodiques pour la caractérisation de locuteurs déterminés ;
- d'un point de vue heuristique, à travers le comptage de valeurs temporelles, et en vue de nous éclairer sur les spécificités rythmiques des locuteurs.

Abordant le champ cognitif, nous nous inscrivons dans la perspective de la « cognition située » (Dubois (dir), 1991 ; Dubois, 2001), qui s'attache à repérer des catégories sémantiques à travers la diversité des pratiques de différents types d'usagers. Nous proposons une réflexion sur la (voire les) catégorisation(s) de la voix par des auditeurs ordinaires.

2.2. Paradigme individualité / parole

Nous intéressant au premier plan, le paradigme individualité et parole, au cœur duquel se situe le sujet parlant, est étudié à travers les notions :

- d'*habitus individuel* pour le premier volet de l'étude ;
- de *références cognitives identifiantes de traits de personnalité* pour le seconde volet.

L'*habitus*, défini au sens de Bourdieu comme « un système de dispositions durables (...), [permet] d'articuler l'individuel et le social, les structures internes de la subjectivité et les structures sociales internes » (Accardo & Corcuff, 1986 : 67-83). Nous retenons alors la notion d'*habitus individuel* ou *subjectif* – opposé à *habitus de classe* – (Bourdieu, 1983), en tant que pratique langagière impliquant un marquage propre à un locuteur donné. Le repérage de marques idiosyncrasiques que nous effectuons à partir de la parole de trois locuteurs se fera ainsi d'abord à travers la recherche d'*habitus prosodiques individuels*.

Sur un plan cognitif, alors que la reconnaissance de visages a donné lieu à des explorations nombreuses dans le champ des recherches cognitives, la reconnaissance des voix a été délaissée, celle de la parole occupant le devant des recherches en particulier pour le développement de systèmes de reconnaissance automatique. Nous abordons donc cette question dans une perspective de la reconnaissance et de l'identification de « sources sonores » conceptualisées dans le cadre des théories de la catégorisation (Dubois, 2000). Dans ce cadre, la voix constitue un indice de quelque chose – ici quelqu'un – en liaison avec les représentations préalablement expérimentées et mémorisées.

À travers les *inférences sémantiques* (Dubois, à paraître), dégagés à partir des séquences verbales suscitées auprès d'auditeurs en réponse à l'écoute des voix des premiers locuteurs, nous explorerons les *références sémantiques et cognitives identitaires* qui en résultent.

3. Méthodologie

3.1. Deux corpus

Un *corpus vocal* est composé d'extraits d'entretiens menés auprès de trois Parisiens âgés (une femme de 75 ans, deux hommes de 88 ans), s'exprimant à propos de leur quartier de résidence, la Butte Montmartre. Nous postulons qu'ils sont garants d'une connaissance passée et présente, collective et individuelle de ce lieu emblématique de Paris, et que le contenu sémantique de leur discours a une certaine homogénéité (des souvenirs du quartier).

Un *corpus commentaires* présente un recueil de verbalisations libres. Celles-ci ont été obtenues à l'issue d'un questionnaire accompagnant un test d'écoute composé d'extraits des voix des locuteurs du *corpus vocal*, qui ont été donnés à écouter à douze sujets (6 femmes et 6 hommes, âgés entre 25 et 50 ans, natifs du français, de milieux socioculturels variés).

3.2. Démarche

Premièrement, nous entreprenons de rendre compte de variations linguistiques observables au plan prosodique, à travers l'analyse des propriétés i) mélodiques et ii) rythmiques spécifiant chaque locuteur du *corpus vocal*.

i) L'application d'un analyseur mélodique¹ vise à repérer des mots *mélismés* dans le discours de chacun des trois locuteurs, *et alors* à spécifier leur usage personnel de ces « ornements mélodiques ». La notion de *mélisme* – issue du vocabulaire du chant – se définissant ici, au sens donné par Caelen-Haumont, comme :

- indice d'un investissement subjectif du locuteur dans son discours,
- combinant les niveaux sémantique, pragmatique et cognitif,
- désignant les niveaux les plus aigus du registre moyen du locuteur,
- repérable automatiquement (avec MELISM, cf. Références).

ii) Le relevé de marques de durée (longueurs et nombres de segments), de débit (nombres de syllabes par seconde), de pauses silence, de « Marques de Travail de Formulation »² (répétitions de mots-outils, *euh...*) a pour but de caractériser leur usage individuel des modalités rythmiques.

Deuxièmement, nous envisageons d'examiner les phénomènes de catégorisation à l'œuvre dans les jugements des auditeurs (*corpus commentaires*), qui ont écouté les extraits de voix. Quelles qualifications les auditeurs sont-ils capables de fournir quant aux propriétés de la voix (qualité sonore), quant aux caractéristiques identitaires ?

Par l'observation de marques linguistiques précises (formes nominales, adjectivales, figures métaphoriques et métonymiques), nous cherchons à discerner les processus langagiers associés à des constructions cognitives distinctives, comme autant de catégories pouvant servir à spécifier les voix des locuteurs nous intéressant.

4. Résultats

4.1. Identification de marques prosodiques individuelles

Notre investigation se rapportant à la mélodie a tout d'abord permis de valider des critères du *mélisme*, notamment :

- sa valeur pragmatique, c'est-à-dire qu'il est fortement lié à la tâche langagière engagée (ici parler de son quartier) et au thème du discours (ici un souvenir de la Butte Montmartre), avec 70% des *mélismes* identifiés qui appartiennent aux champs lexicaux prédominants chez les trois locuteurs (isotopies du souvenir, de lieux et personnages typiques) ;
- sa fonction d'indice d'investissement subjectif du locuteur dans son discours (les *mélismes* apparaissent sur 13,5%, 17% et 13,5% des items chez chacun des locuteurs, donc une part non négligeable de leur discours ; et sont dans les registres mélodiques les plus aigus, comme *artistes*, *chansonnier* et *guerre* par exemple, qui sont employés par la locutrice, et les deux locuteurs respectivement avec un empan mélodique important).

¹ Programme MELISM, implémenté dans l'analyseur PRAAT.

² Morel & Danon-Boileau (1998).

D'autre part, nous avons mis en évidence deux autres phénomènes validant nos propres hypothèses en ce que le *mélisme* est :

- un indicateur discursif efficace car pouvant apparaître sur des unités hétérogènes et différentes de celles se rapportant à des traits sémantiques forts (on a isolé des éléments où se manifeste une forte densité intonative comme des interjections, des extraits chantés, qui se trouvent soit indirectement liés aux champs lexicaux, soit insolites) ;
- surtout un subtil moyen de distinction inter- et intra-locuteurs.

Tout particulièrement, chacun des locuteurs effectue des *mélismes* sur ses champs lexicaux privilégiés (tous sur le thème du quartier Montmartre, objet de discours des entretiens), l'un renvoyant à l'art, le second à l'histoire, le troisième à la topographie. Nous parlons ainsi à ce propos d'*habitus mélodiques individuels*.

4.1.1. Exemple

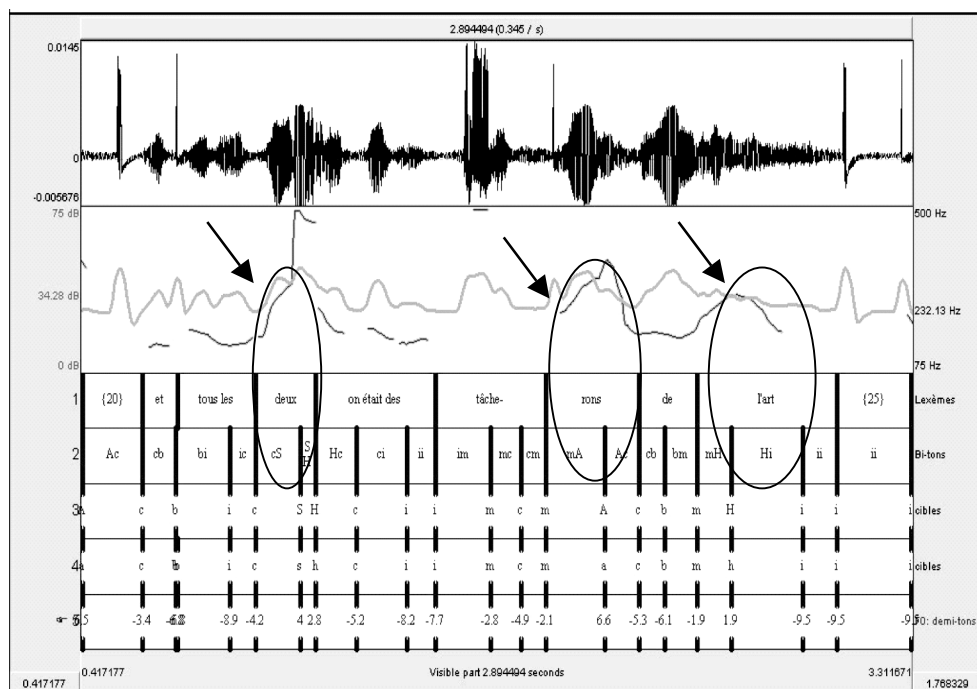


Figure 1 - Spectrogramme (Praat>Melism) illustrant trois mélismes produits par la locutrice 1

Parmi les différents éléments apparaissant sur cette figure, on peut voir trois items saillants (*deux*, *tâcherons*, *art*) (cf. bulles et flèches), grâce à leur fréquence moyenne de large amplitude³, codés en lettres signalant ici les niveaux supra-aigus (S), aigus (A), et haut (H). Ainsi les *mélismes* portant sur *deux*, *tâcherons*, *art*, prononcés par la locutrice 1, renvoient au champ lexical de l'art, caractéristique de son objet de discours.

Ensuite, concernant l'analyse rythmique, des spécificités propres à nos sujets ont pu être distinguées, par-delà une instabilité générale et fluctuante de l'ensemble des données de ce niveau. Nous avons discerné des valeurs d'ordre stylistique :

³ La courbe d'intensité (en Hertz) n'est pas à prendre en compte ici car l'enregistrement est saturé (particulièrement visible sur la courbe correspondant au signal, tout en haut de la figure). Ce défaut n'affecte cependant pas la courbe de fréquence.

- un rythme « travaillé » chez la première locutrice, « pondéré » chez le second locuteur, « accéléré » chez le troisième locuteur ;
- trois styles que nous interprétons respectivement comme une possible stratégie rhétorique, comme exprimant un désir d'intelligibilité, comme une volonté de communiquer.

Nous envisageons ces propriétés comme contribuant à la constitution d'un *habitus rythmique propre* à chacun des sujets.

4.2. *Émergence de catégories qualifiantes de la voix et identifiantes de la personne*

L'étude des commentaires verbaux produits par les auditeurs à l'issue du test de perception⁴, a permis de faire émerger un panel varié de catégories sémantiques et cognitives :

- particulièrement précis dans l'ordonnancement,
- étonnamment consensuel dans l'attribution de traits de personnalité pour chaque locuteur.

La voix est d'abord apparue comme un objet sonore. Celui-ci pouvant être doté de propriétés spécifiques, repérables à travers :

- des figures rhétoriques comme la métaphore (exemple : *voix comme des petites billes de bois*) ou la métonymie (exemple : *voix de radio*) ;
- des adjectifs qualificatifs (exemples : *aigus, grave, éraillée*).

Différemment, cette catégorie de la voix en tant qu'objet sonore peut résulter d'un effet produit sur l'auditeur, manifeste aussi à travers des formes adjectivales (exemples : *agréable, désagréable, sympathique, captivante*).

D'autres éléments lexicaux nous ont conduit à inférer une propriété perceptive distincte, où la voix apparaît décrite comme l'indice d'une personne (Morange, 2004a)⁵. Ceci à travers l'évocation d'une succession de catégories⁶. Ainsi, le *genre*, l'*âge*, sont apparus systématiquement ; la *situation socio-professionnelle*, la *situation géographique*, le *tempérament* et / ou le *comportement en société*, le *rapport à la parole*, la *physionomie*, le *décor domestique* sont apparus différemment selon les auditeurs. Exemples :

- *C'est une femme assez âgée qui parle, je l'imagine avec une permanente bleue.*

<i>genre</i>	<i>âge</i>	<i>physionomie</i>
--------------	------------	--------------------
- *Un homme âgé d'origine provinciale plutôt du nord ou centre (Normandie, Picardie) ou ceinture de Paris, ayant travaillé dans le tertiaire (...)*

<i>Genre</i>	<i>âge</i>	<i>situation géographique</i>
<i>.....</i>		
<i>situation socioprofessionnelle</i>		
<u><i>perte de l'accent,</i></u>	<u><i>vivant en ville,</i></u>	<u><i>calme.</i></u>
<i>rapport à la parole</i>	<i>situation géographique</i>	<i>tempérament</i>

Ces deux axes cognitifs – la voix comme objet sonore en soi, et la voix comme phénomène sonore indiciaire de la personne – permettent alors de faire des inférences sur les propriétés des voix des locuteurs, car les catégories proposées s'établissent à travers un large consensus

⁴ C'est-à-dire l'analyse des réponses qu'ils fournissent par écrit après chaque écoute d'extrait de voix des trois locuteurs du *corpus vocal*.

⁵ Si le contenu des paroles joue pour beaucoup, nous jugeons ici que la voix occupe une grande part de la source des observations des auditeurs, car ces derniers répondent à la question « Que pensez-vous de la voix que vous venez d'entendre ? ».

⁶ Nommées ici sous des vocables illustratifs – *genre, âge, etc.* –, autrement dit à considérer comme des étiquettes, modifiables.

entre les auditeurs. Par exemple, le même locuteur est qualifié par des adjectifs sémantiquement proches, que l’on peut interpréter comme renvoyant à la *musicalité* (avec *voix chantante, claire, gouleyante*), ou à la *constance* (avec *voix calme, posée*).

Somme toute, l’établissement d’un tel répertoire de catégories qualifiantes et descriptives révèle un traitement perceptif non dimensionnel, non uniquement acoustique, mais un traitement davantage sémiotique, investi de catégories sociologiques, cognitives, du domaine de l’expérience individuelle ou collective. En outre, ce recueil de données représente une sorte d’inventaire du répertoire lexical français commun et non expert disponible sur la voix, à tout le moins circonstancié à notre corpus, et concernant la voix parlée.

4.3. Comparaison des deux approches

Enfin, si nous mettons en regard les propriétés identifiées par les sujets lors de la perception, avec les spécificités acoustiques des locuteurs discernées précédemment, nous aboutissons à l’identification et à la caractérisation de sortes de « portraits perceptifs ».

COMPARAISON DES PROPRIÉTÉS		
	DESCRIPTION ACOUSTIQUE (<i>mélismes, rythme</i>)	DESCRIPTION SUBJECTIVE (<i>commentaires du test d'écoute</i>)
Locutrice 1	Mélismes : champs lexicaux renvoyant à l'art et aux individus Rythme : jouant sur le temps de silence comme sur le temps de parole → Habitus interprétés comme un mode injonctif et péremptoire	- femme âgée, Parisienne - milieu artistique mondain dissimulant une origine plus modeste - sûr de sa personne et s'exprimant aisément
Locuteur 2	Mélismes : champs lexicaux renvoyant au temps et à l'Histoire Rythme : débit lent, peu de silence, nombreuses Marques de Travail de Formulation → Habitus interprétés comme un mode prévenant et minutieux	- homme âgé, ouvrier ou bourgeois - de région ou Parisien - aimant évoquer le passé et son patrimoine - expression langagière correcte et style narratif - nostalgique
Locuteur 3	Mélismes : champs lexicaux renvoyant au lieu Rythme : jouant sur le débit rapide et le silence élevé → Habitus interprétés comme un mode de vivacité et de mobilité	- homme ou femme, 50 ou 70 ans - citadin - d'origine populaire - verve langagière, style "titi parisien" - guilleret et vif

Tableau 1 - Comparaison des descriptions acoustiques et subjectives

Comme le présente le tableau ci-dessus, les descriptions acoustiques et perceptives sont nettement congruentes pour chaque locuteur :

- l’image d’une femme mûre, à la personnalité forte sous une réalité plus équivoque, qui a été perçue à la première écoute de la locutrice 1, n’est pas sans lien avec les jeux prosodiques exploités dans les extrêmes et la thématique de l’art ;
- les traits d’un homme âgé, calme, tourné vers le passé, décrit subjectivement au sujet du deuxième locuteur, se retrouvent d’une certaine façon à travers le rythme modéré et le renvoi à l’histoire, présents dans son discours ;
- la vitalité et la volubilité de l’ancien poulbot, brossées à propos du troisième locuteur, corroborent assez fortement le rythme soutenu et la référence à l’espace, appuyés par sa mélodie.

La question n’était pas ici de savoir si ces traits de personnalités sont vrais (réels ?), mais bien de voir qu’ils sont convergents. Aussi, les résultats des deux principales études de cette

recherche se complètent de façon idoine, à travers la conjonction de ces « portraits perceptifs » avec les traits prosodiques individuels.

5. Conclusion

Le phénomène parlé nous a semblé devoir être appréhendé sous l'aspect de l'idiosyncrasie selon plusieurs approches. Les analyses du *mélisme* – comme phénomène acoustique et sémantique – puis de la catégorisation de la voix – comme phénomène acoustique et indice de personne – ont permis de localiser des variations verbo-vocales intra- et inter-individuelles.

Ainsi, nous apportons quelques connaissances en espérant faire saisir à quel point l'individualité est à « fleur de mots et de sons », ne se situant toutefois pas dans des portraits naturalistes, mais dans des variations typiques de la parole à travers des spécificités individuelles. Nous avons ainsi commencé d'en déterminer certaines propriétés acoustiques, mais l'exploration des processus cognitifs qui les construisent tant dans la production que dans la perception reste à approfondir.

Cette réflexion sur le phénomène de la perception du phénomène vocal suscite actuellement de plus larges investigations, axées notamment sur la difficulté que les locuteurs manifestent à mettre en mots leurs expériences sensorielles quelles qu'elles soient, et plus encore lorsqu'il s'agit d'une entité peu lexicalisée dans la langue et duale (objet incarné) comme la voix, parlée ou chantée (Morange *et al.*, 2005b)⁷.

Références

Références bibliographiques

- ACCARDO, Alain & CORCUFF, Philippe (1986). L'habitus ou l'histoire faite corps, in *La sociologie de Bourdieu. Textes choisis et commentés*, Accardo, Alain & Corcuff Philippe (Eds). Bordeaux : Le Mascaret, 67-83
- BOURDIEU, Pierre (1983). Vous avez dit "populaire" ?, *Actes de la Recherche en Sciences Sociales*, 46 : 98-105.
- CAELEN-HAUMONT, Geneviève & BEL, Bernard (2003). Subjectivité et émotion dans la prosodie de parole et du chant : espace, coordonnées et paramètres, in *Perspectives actuelles sur les émotions. Cognition, langage et développement*, Colletta, Jean-Marc & Tcherkassof, Anna (Eds). Sprimont : Mardaga, 141-147.
- CAELEN-HAUMONT, Geneviève (2004). Melodic lexical prominence and melisms: an automatic tonal annotation by INTSMEL and its application to pragmatic analysis, in *Phonologie et phonétique : approches contemporaines*, Durand, Jacques *et al.* Paris : Hermès.
- CASTELLENGO, Michèle & DUBOIS, Danièle (2005). Timbres ou timbres ? Propriété du signal, de l'instrument, ou construction cognitive ?, *Proceedings of the Conference on Interdisciplinary Musicology (CIMO5) Actes du Colloque interdisciplinaire de musicologie*, Montréal (Québec), Canada, 10-12 mars 2005.
- DUBOIS, Danièle (Dir) (1991). *Sémantique et cognition. Catégories, prototypes, typicalité*. Paris : CNRS éd. Sciences du Langage.
- DUBOIS, Danièle (2000). Categories as Acts of Meaning : The Case of Categories in Olfaction and Audition. *Cognitive Science Quarterly*, 1, 35-68.
- DUBOIS, Danièle (2001). Catégorisation, langage et identité : représentations individuelles et constructions symboliques partagées, in *Identité sociale et langage. La construction du sens*, Costalat-Founeau, Anne-Marie (Dir), Paris : L'Harmattan - Logiques sociales, 195-223.

⁷ On peut, sur ce dernier point, rapprocher la voix du timbre musical (Castellengo & Dubois, 2005).

- DUBOIS, Danièle (à paraître). Sens communs et sens commun : expériences sensibles, connaissance(s) ou doxa ?, *Langue Française*.
- MORANGE, Séverine (2003). La musique vocale du sujet parlant. Ou l'expression de la subjectivité et de la variation intra-individuelle à travers la prosodie, in *Le sujet*. Merle, Jean-Marie (coord.), Paris : Ophrys- Bibliothèque de Faits de Langues, 293-306.
- MORANGE, Séverine (2004a). *La voix en tant qu'indice de la personne*, colloque « Norme / hors norme », Journées scientifiques du réseau de sciences cognitives d'Ile de France (RESCIF), 6-7 fév. 2004, Paris, 44.
- MORANGE, Séverine (2004b). *Apport du traitement mélodique automatique dans l'identification du degré d'investissement subjectif et des propriétés individuelles du sujet parlant*, Paris : MIDL2004, coll. Symposium S003, Presses de Telecom Paris, 93-96.
- MORANGE, Séverine (2005a). Approches structurale, prosodique, psycho-cognitive de quelques propriétés linguistiques subjectives de la parole : exemple de trois Montmartrois âgés, *thèse de doctorat*, Morel, Mary-Annick (Dir), Paris 3- Sorbonne Nouvelle.
- MORANGE, Séverine, FONTAINE, Jean-Marc, VOGEL, Corsin, POITEVINEAU, Jacques & DUBOIS, Danièle (2005b). *Appartenances catégorielles d'un objet complexe, la voix : approches linguistique et psycho-cognitive*, Le Sensolier, 3^e Journée, 7 oct., Ivry, CNRS, 14-17.
- MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *Grammaire de l'intonation. L'exemple du français*. Paris : Ophrys - Bibliothèque de Faits de Langues.

Références à des logiciels

- BOERSMA, Paul & WEENINK, David. (Institute of Phonetic Sciences, IFA), Université d'Amsterdam, *Praat* : <http://www.fon.hum.uva.nl/praat/>
- CAELEN-HAUMONT, Geneviève (LPL, Univ. de Provence) & AURAN, Cyril (SILEX, Lille 3), *Intsint* et *Melism* (développés à partir de Momel, défini par HIRST & ESPESSER, 1993).

THE PRODUCTION OF CORONAL CONSONANTS BY FRENCH LEARNERS OF ENGLISH

Stéphane MORTREUX

stephane.mortreux@lpl.univ-aix.fr

*Laboratoire Parole et Langage, Université de Provence, 29 Avenue Robert Schuman
13621 Aix-en-Provence*

Abstract : This study concentrates on the articulatory and acoustic properties of coronal English consonants /t, d, n/ produced by French learners of English. Although /t, d, n/ belong to the phonemic repertoire of both English and French, there are well-established differences in how these consonants are phonetically realized in both languages. Our goal is to determine to what extent such differences can be produced by French native speakers, depending on their experience with English.

1. Introduction

A great deal of attention has been recently devoted to studying the production and perception of speech sounds in a second language. The acquisition of L2 raises a number of major issues, both for fundamental and applied speech research, which relate for example to the neural underpinnings of speech acquisition and to the human brain's possible limited plasticity after puberty. It is well established that the phonetic system of a language is, among the many components of that language, the one which foreign learners find most difficult to acquire. In addition, the accurate production of speech sounds in L2 is often a complex task for the speech motor system, the difficulty depending however on the relationship between the phonological/phonetic systems of L1 and L2.

This paper concentrates on the acoustic and articulatory properties of coronal consonants /t, d, n/ produced by French learners of English. Few studies have been conducted hitherto in that domain, partly because articulatory data in the production of speech are particularly complex to collect and analyse.

Both French and English contain these coronal consonants, but the phonetic realizations of these consonants are significantly different between the two languages: 1- concerning the voicing contrast (e.g. /t/ and /d/) that is realized differently in both languages. French voiced stops are typically prevoiced and thus characterized by a negative Voice Onset Time (VOT), whereas the English ones are not. Furthermore, English voiceless stops in syllable-initial position are aspirated and are characterized by a much longer VOT than French voiceless stops which are not aspirated (Fig. 1); and 2- the configuration and position of the tongue for the realization of these consonants also differ from one language to another (Dart, 1991). The articulatory and acoustic characteristics of English coronals produced by French learners of English have not been studied yet, Dart's study (1991) concentrating only on native productions.

2. Issues addressed

Differences in how /t, d, n/ are produced in English and French are traditionally characterized as follows:

- First, it is well known that whereas phonologically voiced stops, e.g. /d/, are characterized by a shorter VOT than phonologically voiceless stops, e.g. /t/, VOT is overall shorter in French than in English. The question arises of whether VOT in stops will tend to be longer for more advanced French learners of English than for beginners.

- Second, these consonants are said to be lamino-dental in French as opposed to apico-alveolar in English (Dart, 1991), i.e. both the active articulator (tongue blade vs. tip) and the place of articulation (dental vs. alveolar) differ between the two languages. French beginners in English tend to produce English /t, d, n/ as lamino-dental. The issue addressed in this work is whether more advanced French learners will show an articulatory shift towards an apico-alveolar configuration in the production of these consonants.

We regard these differences as subphonemic (lamino-dental and apico-alveolar /t/s are subphonemic differences of the same phoneme), and so one important issue is to what extent French learners can become sensitive to these articulatory differences and learn how to produce them despite their subphonemic nature.

As for the acoustic correlates of the voicing contrast, we expect VOT to have roughly the same duration in both English and French stops produced by beginners in English, as voicing production depends on practice more than on the age of acquisition (Hazan, 1993), whereas VOT should be longer in English than in French stops produced by advanced learners.

3. Corpus and recordings

Two groups of French university students in English were tested. Group 1 was composed of five beginners in English, i.e. French university students in their first year of English and who have never been in an English speaking country (or only for one-week long trips), and Group 2 of five more advanced learners, being either English phonetics teachers at university or Ph.D. students in English phonetics, and having often traveled in an English speaking country. All the speakers are French natives and learned English at school. They have no English relatives, and they all began to study English phonetics (i.e. how to produce the phonemes accurately) at university.

The material was made up of a series of nine phonetically-similar pairs of monosyllabic English and French words or pseudo-words, e.g. *damn* – *dame*. The consonants /t, d, n/, appeared in onset pre-vocalic position in all of the words.

First, an acoustic recording was conducted, each word being spoken within a carrier sentence (in order to create a phonetic context), with five repetitions for each sentence:

- Il reprend ... encore (French words)
- Put up a ... above all (English words)

The different pairs of words used for this study are as follow (English words on the left and French ones on the right):

- tap/tape (strikes) ; tip/type (guy) ; top/top (top) ; damn/dame (lady) ; dab/dab ; deb/deb ; nap/nappe (tablecloth) ; nib/nib ; nip/nippe (to dress)

In all, 900 words were so recorded (450 in English and 450 in French), thus a total of 300 words per consonant.

In a second step, we carried out an articulatory study with the same speakers using static palatography. A mixture of dessert cream and medicinal charcoal was applied on the subject's tongue prior to her/his producing each word in isolation. A digital photography of the palate's image, as reflected by a purpose-built mirror, was taken immediately afterwards and transferred onto a computer for further processing. These images allowed us to characterize the spatial configuration of the tongue-palate contacts in the production of the target consonant.

The pairs tap/tape, damn/dame and nap/nappe have been chosen for analyse (as the initial consonant is followed by an open vowel and thus restricting undesired tongue/palate contacts), with two repetitions for each (the production of /t/ is missing for one advanced learner however), hence a total of 116 words.

4. Results

4.1. Acoustic results

For VOT measures, we based our analyses on Lisker & Abramson's criteria (1964). We thus considered that the VOT is the time interval between the consonant burst and the onset of glottal vibration.

We tested the interaction between Group and Language factors on VOT values through a linear mixed-effects model, with Language (French, English) and Group (Beginners, Advanced learners) as independent variables, and Subjects as random factor. The result proved statistically significant ($t(498) = 3.52701$, $p < 0.01$).

A repeated-measure ANOVA was then conducted, with Consonant (/d/, /t/) and Language (French, English), as independent variables, and subjects as random factor.

The results concerning /t/ and /d/ show (fig. 2) that English stops were produced, as expected, by advanced learners with a longer VOT than French stops, this result proving significant ($F(1,4) = 43.679$, $p < 0.01$). The VOT for English /t/s is even almost twice longer than for French /t/s (respectively 80 ms and 42 ms), and this is also a significant result ($F(1,4) = 21.352$, $p = 0.01$). Also, as regards the production of English /d/s by advanced learners, we can observe a bimodal distribution of the results, some being produced in the French manner (pre-voiced), others being on the contrary devoiced (with a positive VOT). We will discuss this result in the following section.

By contrast, no such difference was found between French and English stops for French beginners, although this result is not significant ($F(1,4) = 1.629$, $p = 0.271$). The VOT values for English and French /t/s are even the same, with an average duration of 46 ms for both languages.

4.2. Palatographic results

In order to classify the photographs into different categories and to decide of the place of articulation, a classification in three groups was made. The productions were considered as dental when the whole back of the incisors was touched by the tongue and thus covered with the mixture. When only part of the back of the incisors was touched and covered, the articulations were classified as dento-alveolars. The third category corresponds to productions which did not touch the incisors at all, being classified as alveolars (see fig. 3 for a sample of each).

As regards English consonants, it appears as expected that the majority are produced as alveolars by advanced learners (69%), and that none of them corresponds to a dental place of articulation. However, but as expected also, most of these consonants are articulated as dentals (60%) when produced by beginners, i.e. correspond to the French consonants properties (fig. 4). Only 2 words out of 30 (7%) were produced "correctly" by beginners, i.e. as alveolars (see table 1 for more details).

Concerning French consonants, it is not surprising to find similar results for advanced learners and beginners, all the speakers being French natives. These consonants are generally produced as dento-alveolars by advanced learners as well as beginners (42% and 40% respectively), but the production also appears to be sometimes dental (especially concerning beginners whose productions are actually divided between dental and dento-alveolar when looking closer at it). More surprisingly, some of these consonants are produced as alveolar for both groups of speakers (see table 2 for full details).

5. Discussion

As regards VOT, we have seen that group 2 (composed of advanced learners) shows the ability to shift the frontier to the right, i.e. producing a longer VOT for English initial stops than for French ones as we expected. However, it can be observed that some of the English /d/s are prevoiced, i.e. have a negative VOT, which is typically French. Even though advanced learners are aware and able to produce a longer VOT for English initial stops than for French ones, it appears that some of the English /d/s are produced as French /d/s, which would mean that the “degree” of advancement is not “perfect”. Concentration must be an important factor here, since these words were inserted into sentences, and so the speakers may have lost their concentration at some times of the recordings. It can be concluded from that point that a greater attention is thus required in order to produce properly L2 phonemes, even though the speaker is very familiar with this L2.

Concerning this bi-modal distribution of the results and to give a possible explanation, the voicing contrast between French and English /d/s concerns the phase relation that settles between laryngeal vibrations and supra-laryngeal articulators movements (opening/closure of the vocal tract). We can suppose, concerning the advanced learners, that the phase relation proper to English /d/ has been acquired, and that the attractor associated with this consonant is present. However, the attractor associated with French /d/ may still be present. The articulatory system would thus have two attractors competing with one another. These results would present interesting implications, particularly concerning the acquisition of an L2.

As for the group of beginners, no difference is shown in their production of English and French initial stops concerning the VOT. In other words, English stops were produced with French VOT values. The argument pointed here does not imply that they are not able to produce these phonemes differently, but that they are either not aware of any difference between English and French phonemes, or if they are aware of any difference in consonants production, they do not know how to produce this difference.

Concerning English consonants and their place of articulation, a shift is observed when produced by advanced learners (as expected). The results show that these speakers are able to shift from a dental (for French consonants) to an alveolar (for English consonants) place of articulation. A study concerning the manner of articulation (in progress) of these consonants (apical in English and laminal in French) may give us more details about this shift observed with advanced learners. Moreover, it may provide us with more evidence that they actually are conscious of the different parts of the tongue used to produce English and French phonemes.

However, no such difference is observed, concerning beginners in English, as it was expected. Here again, their ability to shift the place of articulation is not questioned, we suggest that either they do not hear any difference between French and English consonants and thus produce them similarly, or they actually hear a difference but do not know how to produce it (which means they would be able to produce it if they knew how to articulate these consonants). It would thus be interesting to conduct a perceptive experiment mixing English and French phonemes inside different sentences in both languages in order to see if beginners are actually able to differentiate French and English consonants.

6. Conclusion

This study concentrating on the production of L2 phonemes /t, d, n/ by L2 learners shows that the speakers are able to produce accurate L2 sounds, even if they have been learned after puberty. The advanced learners participating in this experiment learned English at school, and all began to study English phonetics (i.e. began to learn how to produce L2 sounds properly) at university. This is particularly interesting concerning the plasticity of the brain

(which allows the speakers to produce L2 consonants accurately), and also the question of the age of acquisition. It is generally argued that L2 phonemes have to be acquired very early to become native-like. Flege (1987) even concluded from his experiment on French learners of English and on English learners of French, and concerning the VOT, that it was impossible to produce accurately L2 phonemes that also exist in L1 (what he calls “similar” phonemes). Yet we have seen, particularly for English /t/s, that they were pronounced by advanced learners with twice longer VOT values than French /t/s. However, Hazan & al. (1993) concluded that the voicing contrast, between French and English, and its accurate realization by French-English bilinguals, depend more on practice than on the age of acquisition, which seems to be confirmed with our present study.

We pointed out in our introduction that the accurate production of speech sounds in L2 is a complex task for the speech motor system. From this study, we can argue that the advanced learners are sensitive to the articulatory differences between L1 and L2 phonemes, despite the fact that these differences are subphonemic and thus represent a complex task in their realization for the speech motor system. As the speakers are able to produce differently L1 and L2 phonemes presenting very subtle differences in their realization, it would be interesting to conduct a perceptive experiment in order to analyse if this sensitivity to subphonemic differences exists not only in the production of these sounds, but also in their perception.

7. Acknowledgements

We want to thank Bernard Teston who managed to build the system needed to conduct the palatagraphic experiments. Without his help this experiment could not have been conducted. We also wish to warmly thank all the speakers who kindly agreed to participate in this experiment. This work being part of our Ph.D., we want to thank Noël Nguyen for his supervision and his always valuable advice in the writing of the article.

8. Figures and tables

8.1. Tables

	Dental	Dento-alveolar	Alveolar
Advanced learners	0%	31%	69%
Beginners	60%	33%	7%

Table 1 - Place of articulation of English /t, d, n/ by advanced learners and beginners.

	Dental	Dento-alveolar	Alveolar
Advanced learners	31%	42%	27%
Beginners	37%	40%	23%

Table 2 - Place of articulation of French /t, d, n/ by advanced learners and beginners.

8.2. Figures

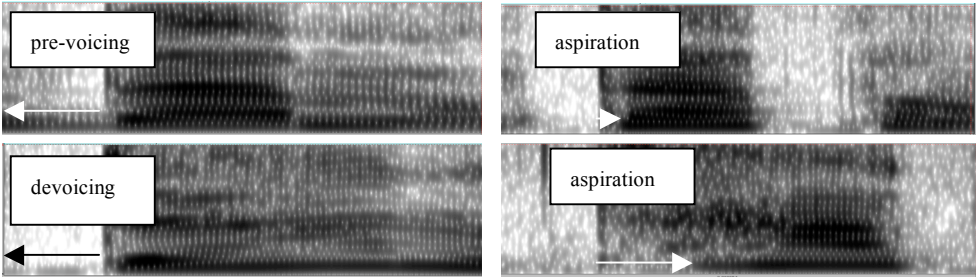


Figure 1 - Spectrograms (samples from our corpus) of French and English /d/s (*dame-damn*) on the left and /t/s (*tape-tap*) on the right, with French consonants on top and English ones at the bottom.

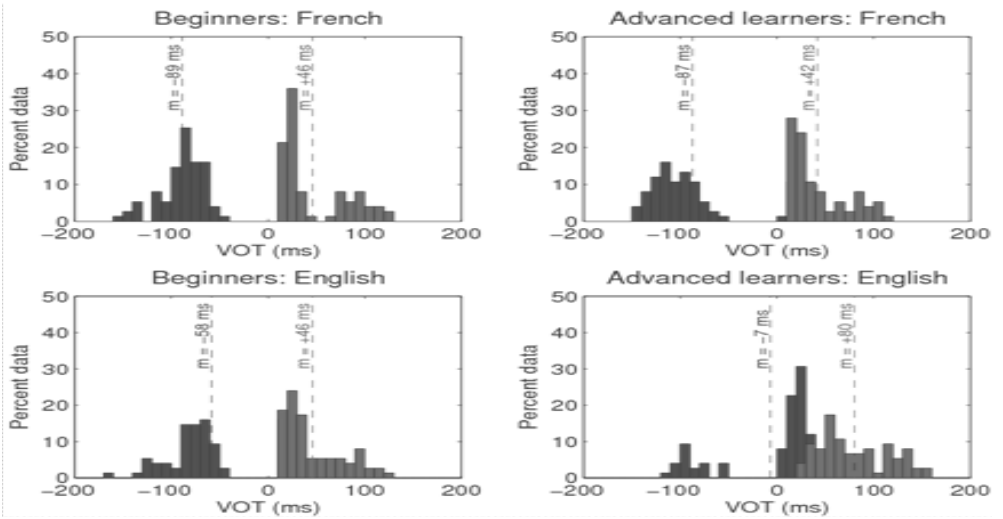


Figure 2 - VOT values for /t/s (grey) and /d/s (black) produced by beginners and advanced learners for English and French consonants.

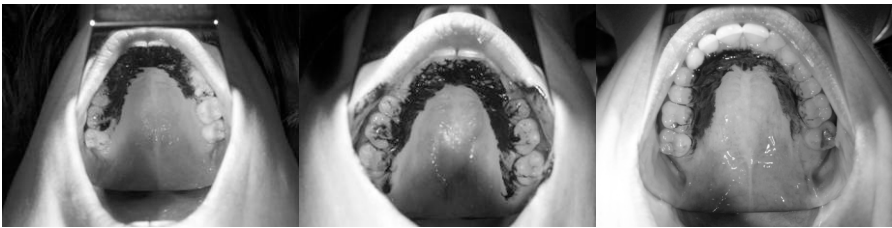


Figure 3 - Articulatory classification in 3 categories: dental (left), dento-alveolar (middle), and alveolar (right).



Figure 4 - English /d/ (damn), /n/ (nap), /t/ (tap) from left to right, produced by an advanced learner (top) and a beginner (bottom).

References

Bibliographic references

- DART, Sarah (1991). *Articulatory and Acoustic Properties of Apical and Laminal Articulations*. UCLA Working Papers in Phonetics, 79.
- FLEGE, James Emil (1987). The production of “new” and “similar” phones in a foreign language: evidence for the effect of equivalence classification, *Journal of Phonetics*, 15: 47-65.
- HAZAN, Valérie, BOULAKIA, Georges (1993). Perception and production of a voicing contrast by French-English bilinguals, *Language and Speech*, 36: 17-38.
- LISKER, Leigh, ABRAMSON, Arthur (1964). A cross-language study of voicing in initial stops: Acoustical measurements, *Word*, 20: 384-422.

ESSAI DE DESCRIPTION D'EMPLOIS DES MOTS DU DISCOURS AUTOUR DE *HEIN*

Hiroko NODA
hrknd@yahoo.fr
Laseldi-Idiomes
 30 rue Mégevand
 25030 Besançon Cedex

Abstract : The aim of this paper is to show the importance of discourse markers in the organisation of discourse. In particular, it analyses the syntactic, semantic, pragmatic and prosodic functions of *hein* in French, according to the "Theory of Predicative and Enunciative Operations" developed by Culioli.

1. Introduction

Comment peut-on se rendre compte des emplois des locutions comme *ben*, *dis*, *écoute*, *hein*, *n'est-ce pas*, *quoi*, *tiens*, *tu sais*, etc.¹ dans un discours oral chez les natifs français ? Mon objectif de recherche est de décrire en particulier des emplois de *hein* en les comparant avec ceux de *n'est-ce pas* qui est souvent traité comme synonyme² et de *quoi* dont les similitudes ont été mentionnées du point de vue fonctionnel ou prosodique³.

Dans les recherches linguistiques, les mots du discours (MD) qui ont souvent été considérés comme « ratés » ou « parasites » du discours oral, n'ont pas été l'objet principal d'analyses jusqu'à la fin des années soixante, comme le remarque Settekorn (1977 : 195). Depuis, de nombreuses recherches ont traité les emplois des MD⁴. Mais les descriptions des MD ont souvent la forme d'une liste de fonctions pragmatiques qui ne permet pas forcément d'appréhender la totalité de leurs propriétés syntaxiques, sémantiques, pragmatiques et prosodiques et de différencier les emplois de chaque MD.

En m'appuyant sur la « Théorie des Opérations Prédicatives et Énonciatives » développée par Culioli, je tenterai de faire une description des emplois de *hein*, en essayant de faire apparaître la différence avec ceux de *n'est-ce pas* et de *quoi*, afin de montrer l'importance du rôle qu'ils jouent dans l'organisation de la parole.

2. Méthodologie

Pour atteindre cet objectif, j'ai choisi de travailler sur un corpus constitué de séquences tirées des émissions de télévision « À l'école des palaces » du 27 février 2006 (noté « epl ») et « Ça se discute, jour après jour » du 5 avril 2004 (noté « csd »), des interviews « la fleuriste » (noté « fle ») et « l'infirmière » (noté « inf ») qui sont tirés du « Choix de textes de français parlé » (Blanche-Benveniste *et al.*, 2002), et du film « Chacun cherche son chat » (noté « ccc ») de Cédric Klapisch. Ce choix est dû au fait que ces séquences comportent un

¹ Pour désigner cet ensemble d'expressions, j'ai repris l'appellation « mots du discours » qui a été introduite par Ducrot *et al.* (1980) et qui me paraît plus générale et plus neutre. Je n'entrerai pas ici dans le problème des définitions et des appellations qui sont différentes suivant les auteurs.

² Par exemple, Andrews (1989).

³ Entre autres, Morel & Danon-Boileau (1998), Beeching (2004).

⁴ Auchlin (1981), Gülich & Kotschi (1983), Luzzati (1985), Vincent (1993), Fernandez (1994), Mosegaard Hansen (1998), Morel & Danon-Boileau (1998), Beeching (2002), Chanet (2004), Dostie (2004), etc.

langage familier et populaire employé dans une situation quasiment naturelle. Les exemples cités sont transcrits en suivant en général les conventions fixées par le centre de recherche VALIBEL (Dister *et al.*, 2006)⁵. Elles consistent à respecter au maximum l’orthographe standard sans ponctuation ni majuscule sauf pour les noms propres. Pour les éléments prosodiques, j’ai utilisé le programme Praat⁶ et le module « Prosogram » développé par Piet Mertens qui permet de faire une stylisation des variations de hauteur de la voix en prenant en compte la perception tonale.

3. Hypothèses

J’aimerais présenter les hypothèses sur les emplois de *hein*, de *n’est-ce pas* et de *quoi* élaborées dans le cadre de ma thèse en cours. Ils s’inscrivent tout d’abord dans une relation intersubjective énonciateur (noté S₀) / coénonciateur (noté S’₀). Admettons que cette relation intersubjective soit constituée de la notion prédicative P qui peut prendre une valeur positive *p* ou une valeur négative *p*’ (autre valeur que *p*). Admettons également que S₀ comme instance subjective à l’origine d’un énoncé soit garant de la validité / validabilité de cet énoncé et que S’₀ comme instance subjective séparable (mais non nécessairement séparée) de S₀ soit un pôle d’altérité possible sur l’énonciation de S₀.

Hein met en place S’₀ comme covalideur de la valeur *p* choisie par S₀ afin de renforcer sa prise de position et non pas de vérifier cette validité. Dans le cas de *n’est-ce pas*, S₀ qui a choisi la valeur *p* demande à S’₀ de faire une validation de cette valeur. La valeur *p* est pondérée, c’est-à-dire que S₀ attend que S’₀ choisisse la valeur *p*, mais la valeur *p*’ n’est pas complètement exclue. *Quoi* marque que S₀ considère qu’il est évident que S’₀ choisira la valeur *p*.

4. Résultats

Je vais essayer de mettre en œuvre ces hypothèses sur les emplois de *hein*, de *n’est-ce pas* et de *quoi*⁷ en citant quelques exemples significatifs qui marquent des contraintes distributionnelles et contextuelles. En comparant les énoncés suivants, il apparaît que *hein*, *n’est-ce pas* et *quoi* (1a, 1b, 1c) s’inscrivent dans le cadre d’une relation intersubjective S₀ / S’₀, à la différence de l’énoncé (1d) qui ne se situe pas forcément dans cette relation :

- (1a) enfin il était très jaloux hein (csd)
- (1b) enfin il était très jaloux n’est-ce pas
- (1c) enfin il était très jaloux quoi
- (1d) enfin il était très jaloux

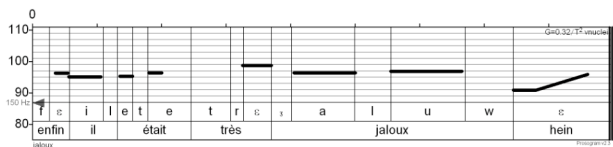


Figure 1 - Prosogram de l'exemple (1a)⁸

Quelles sont alors les différences entre *hein*, *n’est-ce pas* et *quoi* ? Cet emploi de *hein* est répertorié souvent comme « demande de consentement » dans les études précédentes. Dans

⁵ Ces conventions sont largement convergentes avec celles de la DELIC ou de la PFC (Delais-Roussarie *et al.*, 2002).

⁶ Praat est un outil d’analyses phonétiques développé par Paul Boersma et David Weenink.

⁷ Vu le nombre de pages limité pour cet article, je ne travaille que sur les emplois de *quoi* qui sont comparables avec ceux de *hein*.

⁸ La dimension de l’image dans cet article est appropriée à la longueur de chaque énoncé.

la situation originale, la locutrice parle de son ex-mari au journaliste qui n'est pas censé le connaître. Or cet énoncé peut être également adressé à quelqu'un qui le connaît. C'est-à-dire qu'il est possible d'employer *hein* indépendamment de la capacité de l'interlocuteur de juger la véracité du propos. En effet, le locuteur ne fait pas une véritable demande d'approbation dans la mesure où il considère que l'interlocuteur est de son avis. L'énoncé reste une demande mais qui n'attend qu'une réponse positive. *Hein* marque que S'_0 est mis en place comme instance qui valide la valeur choisie par S_0 , c'est-à-dire comme covalideur.

Du point de vue prosodique, on peut observer une mélodie montante de *hein* après une accentuation sur l'élément précédent (ici *jaloux*) (Figure 1). Ce type de mélodie est fréquent dans le cas de *hein* en fin d'énoncé comme l'exemple (1a).

Dans le cas de *n'est-ce pas*, l'énoncé est plus interrogatif que dans le cas de *hein* dans le sens où le locuteur n'exclut pas la réponse négative de la part de l'interlocuteur, même si la réponse attendue tend à être positive. Le locuteur demande si l'interlocuteur est d'accord avec lui, même s'il ne pose pas une vraie question, c'est-à-dire celle de savoir si l'interlocuteur est de son avis. Dans ce cas, S_0 qui a choisi la valeur p demande à S'_0 de faire une validation de cette valeur. La valeur p est pondérée mais la valeur p' n'est pas totalement exclue, ce qui est différent des emplois de *hein* où la valeur p' est exclue.

L'exemple avec *quoi* (1c) ressemble plus à une affirmation par rapport aux exemples avec *hein* (1a) et *n'est-ce pas* (1b), mais par rapport à l'exemple (1d), *quoi* comporte un aspect plus intersubjectif. Dans le cas de *quoi*, le locuteur qui exprime son avis comme une évidence ne laisse pas le choix à l'interlocuteur d'exprimer son opinion. S_0 qui a choisi la valeur p ne demande pas une validation de cette valeur à S'_0 à la différence du cas de *n'est-ce pas*. S'_0 est mis en place comme instance qui valide la valeur choisie par S_0 mais à la différence de *hein*, S_0 considère qu'il est évident que S'_0 choisira la valeur p .

L'exemple (2) montre qu'il est difficile d'employer *n'est-ce pas* avec un prédicat subjectif comme « j'ai faim » dont le S'_0 ne peut être garant, contrairement à *hein* et *quoi*. Il est difficile de faire une validation sur un procès dont on ne peut juger la véracité. Cela valide l'hypothèse selon laquelle *n'est-ce pas* implique plus d'interrogativité que *hein* et *quoi* et que *n'est-ce pas* marque une demande de validation.

- (2a) je commence j'ai faim *hein* (csd)
 (2b) ?? je commence j'ai faim *n'est-ce pas*
 (2c) je commence j'ai faim *quoi*

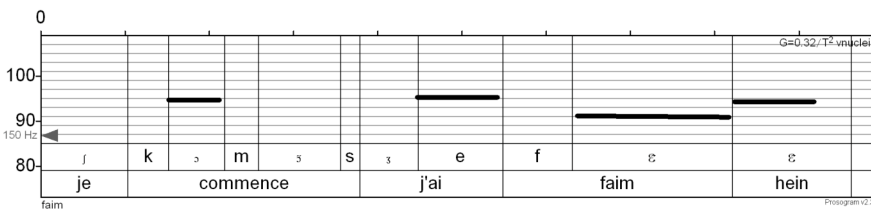


Figure 2 - Prosogram de l'exemple (2a)

L'emploi de *quoi* est possible, mais il donne l'impression que le locuteur est indifférent à ce que pensent les autres, comme s'il se justifiait et leur imposait son action, alors qu'avec *hein*, le locuteur manifeste plus d'attention aux autres, comme s'il demandait une permission. Le statut de S'_0 en tant que covalideur de la valeur choisie par S_0 implique plus d'évidence dans le cas de *quoi* que dans le cas de *hein*.

Dans l'aspect prosodique, il est possible d'observer qu'il y a une descente mélodique avant *hein* et une montée qui porte sur *hein* (Figure 2). À la différence de l'exemple (1a), la mélodie de *hein* est plate mais elle est plus haute que celle qui la précède. D'ailleurs,

l'énoncé (1a) peut être reproduit avec ce type de mélodie. Dans ce cas, l'énoncé impliquerait moins d'interactivité, c'est-à-dire que le locuteur affirme plus son avis et réduit la possibilité d'intervention de l'interlocuteur.

Passons aux exemples (3) et (4) où *hein* est employé après une pause⁹ qui suit un énoncé auquel *hein* est lié du point de vue sémantique. L'exemple (3) montre qu'il est difficile d'employer *n'est-ce pas* et *quoi* à la place de *hein* en gardant la pause.

- (3a) *m'en veux pas* /¹⁰ *hein* (ccc)
- (3b) ?? *m'en veux pas* / *n'est-ce pas*
- (3c) ?? *m'en veux pas* / *quoi*

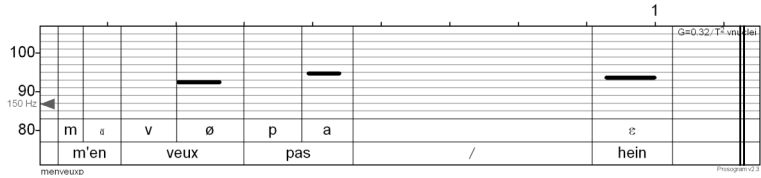


Figure 3 - Prosogram de l'exemple (3a)

Dans ce cas, *hein* sert à renforcer la demande « m'en veux pas » comme s'il la reprenait. Il peut être glosé comme « tu es d'accord ? », « tu es sûr ? ». *Hein* marque que S_0 suppose que S'_0 valide la valeur p validée par S_0 . En revanche, *n'est-ce pas* ne fonctionne pas comme une reprise de la demande qui ne permet pas à S'_0 de valider une autre valeur que celle qui est validée par S_0 . Avec *n'est-ce pas*, la possibilité que S'_0 valide une autre valeur que S_0 n'est pas exclue. S'il n'y a pas de pause qui le précède, *n'est-ce pas* pourrait être employé mais dans ce cas « m'en veux pas » ne fonctionne plus comme une demande de ne pas en vouloir au locuteur. « M'en veux pas n'est-ce pas » serait plutôt une vérification du fait que l'interlocuteur n'en veut pas au locuteur. S'il n'y a pas de pause, *quoi* pourrait fonctionner comme dans l'exemple (2) où S_0 impose sa position à S'_0 . Mais dans le cas où *quoi* est isolé de l'énoncé précédent par une pause, il est difficile d'employer *quoi*, comme le montre également l'exemple (4) :

- (4a) *on va se faire belle hein Mali* /¹¹ *hein* (csd)
- (4b) *on va se faire belle hein Mali* // *n'est-ce pas*
- (4c) ?? *on va se faire belle hein Mali* // *quoi*

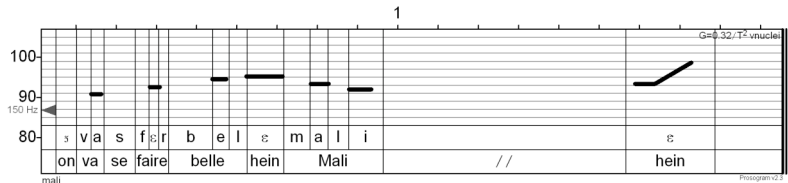


Figure 4 - Prosogram de l'exemple (4a)

Dans ce cas, même s'il n'y a pas de pause, l'emploi de *quoi* est difficile à moins qu'on enlève également *hein* et *Mali*. Cette incompatibilité avec *hein* et l'appellatif *Mali* montre

⁹ Le seuil de pause silencieuse utilisé dans la plupart des recherches est 0,20 ou 0,25 secondes (Cf. Demers (1998 : 63), Campione (2004 : 189), etc.). J'ai appliqué celui de 0,25 secondes dans cet article.

¹⁰ Le symbole « / » marque une pause courte.

¹¹ Le symbole « // » marque une pause plus longue que celle qui est notée par « / ».

que *quoi* ne porte pas forcément une fonction interactive. Avec *quoi*, même si la présence de S'_0 est prise en compte, S_0 ne lui laisse pas le choix d'accepter ou non sa position.

L'usage de *hein* dans l'exemple (4) est proche de celui de l'exemple (3) du point de vue syntaxique et sémantique. Le locuteur demande si l'interlocuteur est d'accord avec sa proposition, en supposant qu'il est de son avis. Du point de vue prosodique, *hein* est marqué par une mélodie montante (Figure 4) à la différence de l'exemple (3) où la mélodie de *hein* est plate (Figure 3). Dans l'exemple (4), il est possible d'employer *hein* avec une mélodie plate, mais cela implique moins de volonté de la part du locuteur de susciter une réaction.

Dans cet exemple, *n'est-ce pas* peut être glosé comme « je pense que tu es d'accord avec ma proposition, est-ce bien le cas ? ». En supposant que S'_0 soit d'accord avec lui, S_0 le vérifie, contrairement au cas de *hein* où l'accord de S'_0 est supposé et n'est pas remis en question.

Comme le montre l'exemple (5), *hein* peut apparaître au sein d'un énoncé.

(5) *alors là c'est difficile parce que bon ils arrivent toujours / plus ou moins / bon pratiquement tous en fait hein // avec euh une séropositivité puis même des sidas déclarés (inf)*

Dans l'exemple (5), *hein* est lié à l'élément adverbial dans une démarche de précision. C'est-à-dire qu'il y a d'abord *toujours* qui apporte une modalité à l'énoncé « ils arrivent avec une séropositivité » et ensuite une modalisation de *toujours* avec *plus ou moins*, *pratiquement tous* à laquelle *hein* ajoute une modalité énonciative. Ces deux niveaux de modalisation sont marqués également au niveau prosodique, par une pause après *hein*, même s'il y a une continuité sémantique. Dans ce cas, l'infirmière explique la situation de ses patients à l'interviewer qui est extérieur à cet hôpital. La locutrice n'attend pas de réaction de l'interlocuteur mais sa présence est prise en compte.

L'exemple (6) peut être considéré comme un cas où *hein* se trouve en incise de la parole.

(6) *un bonbon ça me plaisait pas trop hein j'aimais mieux de l'argent (fle)*

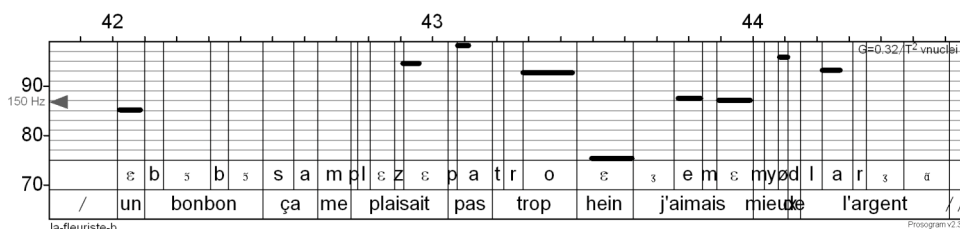


Figure 5 - Prosogram de l'exemple (6)

On peut observer une coupure avant et après *hein* au niveau syntaxique, mais il n'y a pas de pause entre les deux énoncés. *Hein* est marqué avec une hauteur très basse par rapport à son entourage (Figure 5), comme s'il ponctuait la parole. Dans ce cas, la fleuriste en retraite raconte ses souvenirs du travail à l'interviewer. Il n'y a pas de forte volonté interactive dans la mesure où la locutrice n'attend pas forcément une réponse verbale. Mais *hein* marquerait une certaine intersubjectivité, comme si la locutrice essayait d'aider l'interlocuteur à suivre son histoire.

Citons un exemple où *hein* et *quoi* (le deuxième dans l'exemple (7)) peuvent avoir des mélodies similaires : une descente après une montée des phonèmes précédents (Figure 6).

(7) *dans le quand tu es quand tu es dans le speed du service ou quoi c'est un peu casse-pieds quoi pour toi et pour eux hein parce que : (epl)*

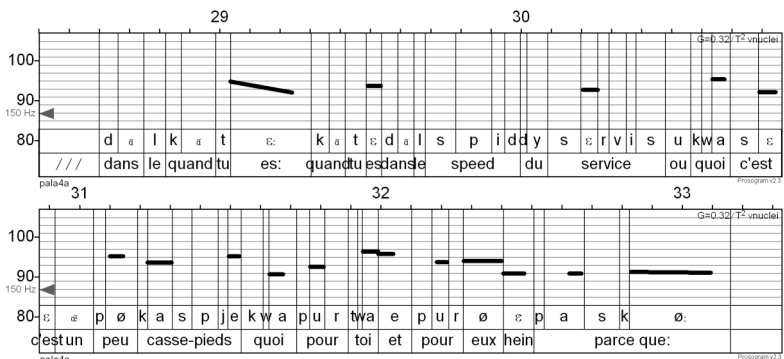


Figure 6 - Prosogram de l'exemple (7)¹²

Ces *hein* et *quoi* sont proches également du point de vue sémantique. La locutrice présente son jugement sur la situation de l'interlocuteur. Par rapport à « c'est casse-pieds », « pour toi et pour eux » – qui porte en partie sur l'interlocuteur – comporte plus d'intersubjectivité. Le fait que *hein* suit un énoncé plus intersubjectif que celui qui suit *quoi* illustre la différence de leurs emplois. Les deux marquent une forte prise de position de S₀, mais dans le cas de *hein*, S₀ demande à S'₀ de valider sa position à la différence du cas de *quoi* où S₀ considère évident que S'₀ la validera.

Quoi peut être marqué par d'autres mélodies, comme le montrent les exemples (8) et (9).

(8) *ça m'a refroidi un petit peu ça m'a: bon sur le coup: euh voilà: quoi moi j'étais un peu surpris: quoi parce que par téléphone c'était pas ça du tout: quoi (epl)*

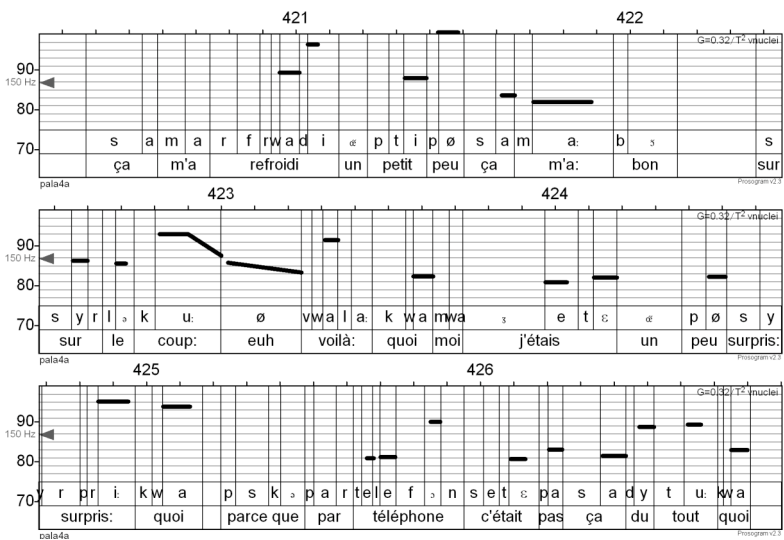


Figure 7 - Prosogram de l'exemple (8)

Dans cet exemple, le premier *quoi* et le troisième *quoi* ont une courbe mélodique similaire : elle est plus basse par rapport à celle qui la précède (Figure 7). Ces emplois de *quoi*

¹² Quand il comporte une césure, le mot est repris au début de la ligne suivante. Par exemple, *c'est* apparaît deux fois (à la fin de la première ligne et au début de la deuxième ligne) dans la figure 6, mais il n'est prononcé qu'une fois.

marquent que S_0 présente son point de vue comme s'il était évident également pour S'_0 . Dans le cas du deuxième *quoi*, la hauteur de la voix ne descend pas autant que pour les deux autres et reste haute. Cette mélodie haute renforce le sens de *quoi* avec lequel le locuteur présente son sentiment comme important et évident pour lui et pour l'interlocuteur.

L'exemple (9) montre un autre type de mélodie pour *quoi* :

(9) *et quand on sort pas la semaine pour un emploi qu'on n'apprécie pas alors là c'est dur quoi (epl)*

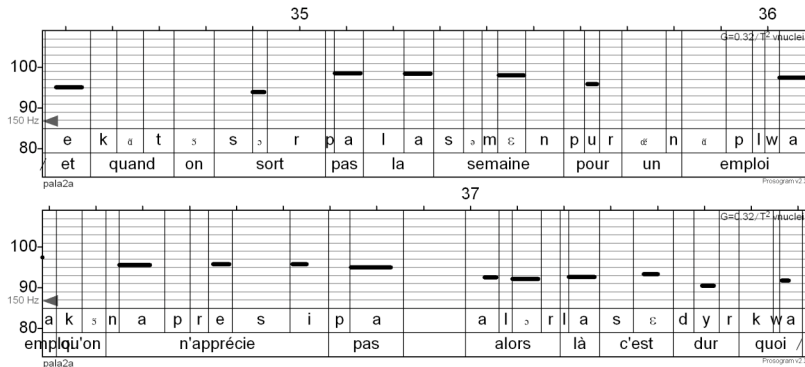


Figure 8 - Prosogram de l'exemple (9)

Dans ce cas, la mélodie de *quoi* est dans la continuité de ce qui la précède « *dur* » (Figure 8). Le trait est relativement court et reste au même niveau. Il n'y a pas de montée mélodique qui la précède comme dans l'exemple (8). Ces caractéristiques prosodiques rendent plus catégoriques ce *quoi* qui marque que le locuteur présente un avis indiscutable et évident.

5. Conclusion

Par rapport à *hein* et *quoi*, *n'est-ce pas* reste interrogatif dans la mesure où S_0 demande si sa valeur choisie est valable pour S'_0 , même si la réponse positive est présumée : « *Je pense que tu es de mon avis, est-ce bien le cas ?* ». En revanche, dans le cas de *hein* et de *quoi*, S_0 considère que S'_0 valide la valeur choisie par S_0 . Avec *quoi*, le fait que S'_0 valide cette valeur implique une évidence, autrement dit, S'_0 n'a pas d'autre choix, alors que dans le cas de *hein*, S_0 demande à S'_0 de valider la valeur choisie par S_0 en attendant une réponse positive.

Les éléments prosodiques ajoutent une grande complexité à leurs emplois. Ils influencent par exemple le degré d'appel d'implication de S'_0 . Il est vraisemblable que la mélodie montante marque que S_0 demande plus d'implication de S'_0 .

Il manque encore des exemples pour affiner ces analyses, mais il apparaît néanmoins que ces marqueurs jouent un rôle non négligeable dans l'intersubjectivité de la parole.

Références

Références bibliographiques

- ANDREWS, Barry J. (1989). Terminating devices in spoken French, *IRAL*, 27 : 193-216.
 AUCLIN, Antoine (1981). Réflexions sur les marqueurs de structuration de la conversation, *Études de linguistique appliquée*, 44 : 88-103.
 BEECHING, Kate (2002). *Gender, Politeness and Pragmatic Particles in French*. Amsterdam/Philadelphia : John Benjamins.
 BEECHING, Kate (2004). Pragmatic particles – polite but powerless? Tone-group terminal *hein* and *quoi* in contemporary spoken French, *Multilingua*, 23: 61-84.

- BLANCHE-BENVENISTE, Claire, ROUGET, Christine & SABIO, Frédéric (2002). *Choix de textes de français parlé*. Paris : Honoré Champion.
- CAMPIONE, Estelle (2004). Étude des interactions entre pauses silencieuses et pauses remplies en français parlé, *Recherches sur le français parlé*, 18 : 185-200.
- CHANET, Catherine (2004). Fréquence des marqueurs discursifs en français parlé : quelques problèmes de méthodologie, *Recherches sur le français parlé*, 18 : 83-107.
- CULIOLI, Antoine (1990-99). *Pour une linguistique de l'énonciation 1-3*. Paris : Ophrys.
- DEMERS Monique (1998). *La prosodie du discours rapporté*. Québec : CIRAL.
- DOSTIE, Gaétane (2004). *Pragmaticalisation et marqueurs discursifs : analyse sémantique et traitement lexicographique*. Bruxelles : De Boeck et Larcier, Duculot.
- DUCROT, Oswald, BOURCIER, Danièle, BRUXELLES, Sylvie, DILLER, Anne-Marie, FOUQUIER, Éric, GOUAZÉ, Jean, MAURY, Luc, NGUYEN, Thanh binh, NUNES, Geraldo, RAGUNET DE SAINT-ALBAN, Laurence, RÉMIS, Anna & SIRDAR-ISKANDAR, Christine (1980). *Les mots du discours*. Paris : Minuit.
- FERNANDEZ, M. M. Jocelyne (1994). *Les particules énonciatives*. Paris : PUF.
- GÜLICH, Elisabeth & KOTSCHI, Thomas (1983). Les marqueurs de la reformulation paraphrastique, *Cahiers de linguistique française*, 5 : 305-351.
- MOSEGAARD HANSEN, Maj-Britt (1998). *The function of discourse particles*. Amsterdam/Philadelphia : John Benjamins.
- LUZZATI, Daniel (1985). Analyse périodique du discours, *Langue Française*, 65 : 62-73.
- MOREL, Mary-Annick & DANON-BOILEAU, Laurent (1998). *Grammaire de l'intonation*. Paris : Ophrys.
- SETTEKORN, Wolfgang (1977). Pragmatique et rhétorique discursive, *Journal of Pragmatics*, 1 : 195-210.
- VINCENT, Diane (1993). *Les ponctuels de la langue et autres mots du discours*. Québec : Nuit Blanche.

Références aux sites Internet

- BOERSMA, Paul, WEENINK, David. Praat, <http://www.praat.org>, (nd). Consulté en février 2007
- DELIC. Corpus du français parlé, <http://www.up.univ-mrs.fr/delic/corpus/index.html>, (nd). Consulté en février 2007
- DELAIS-ROUSSARIE, Elisabeth, DURAND, Jacques, LYCHE, Chantal, MEQQORI, Abderrahim & TARRIER, Jean-Michel. Transcription des données : Outil et conventions, http://www.projet-pfc.net/?u_s=4&u_a=108&sid=, 15/04/2002. Consulté en février 2007
- DISTER, Anne, FRANCARD, Michel, GERON, Geneviève, GIROUL, Vincent, HAMBYE, Philippe, SIMON, Anne Catherine & WILMET, Régine. Conventions de transcription régissant les corpus de la banque de données VALIBEL, <http://valibel.fltr.ucl.ac.be/>, 10/03/2006. Consulté en février 2007
- MERTENS, Piet. The Prosogram, <http://bach.arts.kuleuven.be/pmertens/prosogram/>, 08/02/2005. Consulté en février 2007

LAZY AGREEMENT IN THE LADIN DP

Sabrina RASOM

sabry.rasom@tin.it

Department of Linguistics - University of Padua

35137 Padova

Italy

Abstract : In this article I will describe lazy agreement in the feminine plural noun phrase as it manifests itself in some varieties of Central Ladin (a Romance linguistic domain of Northern Italy). I will base my assumptions on Cinque's theory (2005) about the dual source of adjectives in Romance languages and their different interpretations. I will try to demonstrate how the morphology of lazy agreement disambiguates the possible readings of Romance postnominal adjectives.

1. Introduction

This paper focuses on a morphosyntactic phenomenon characteristic of some Central Ladin varieties. Central Ladin is a Romance linguistic domain extending from Trentino Alto-Adige to Veneto¹; it includes the following varieties: *Fassan*, *Gherdener*, *Badiot*, *Ampezan* and *Fodom*. Fassan Ladin has three further subvarieties: *cazet*, *brach* and *moenat*.

Lazy agreement is a particular phenomenon of agreement in the feminine plural noun phrase (NP) present only in the following Central-Ladin varieties (Rasom 2003, 2005): *Fassan-cazet*, *Gherdener* and *Ampezan*. Basing my assumptions on Cinque (2005) and more particularly on his theory about the dual source of adjectives and their different interpretation in Romance languages, I will describe and analyse this particular construction. My attempt will be to demonstrate how, in these varieties, morphology disambiguates the possible interpretations of postnominal adjectives.

However before this demonstration, I will give a short overview of the phenomenon of the Ladin Lazy Agreement Rule (LLAR), as it was defined by Haiman-Benincà (1992) and I will present the first generalisations deriving from the study of this rule in *Fassan Ladin-cazet*, my native variety (section 2). Moreover, I will analyse the structure of the Romance determiner phrase (DP) and the main generalisations in Cinque's work with regards to the different interpretations of prenominal and postnominal adjectives and their dual source (section 3). I will conclude by interpreting the LLAR in light of Cinque's theory (section 4).

¹ Trentino Alto-Adige and Veneto are two regions of northern Italy.

2. Lazy agreement in the Ladin DP

According to the LLAR, feminine plural NPs receive the mark of number *-(e)s* only on the last member of the phrase, while all other members are only marked with the feature of gender. Three different constructions can figure in this context:

construction:			interpretation:
(1)	la piccola cèses		"the houses, which are all small"
	the small[+f. -pl.] houses[+f. +pl.]		
(2)	la cèses picole		"the houses, which are all small"
	the houses[+f. +pl.] small[+f. +pl.]		
(3)	la cèsa picole		"those houses which are small"
	the house[+f. -pl.] small[+f. +pl.]		

Table 1 - The different interpretations of the Ladin Lazy Agreement Rule according to the different constructions

The interpretation of (2) corresponds exactly to that of (1). The fourth logical possibility is excluded in all interpretations in the Central Ladin varieties with LLAR, while it is possible in the four other varieties (*badiot*, *fodom*, *brach*, *moenat*) in which there is no LLAR, as shown in the examples (4) and (5).

The difference between the examples (4) and (5) is that in three of these varieties (*brach*, *moenat*, *fodom*) there is no sigmatic plural (ex. 5):

- (4.) (*) les picole cèses
the small[+f.+pl.] houses[+f.+pl.]
- (5.) (*) le picole ciase
the small[+f.+pl.] houses[+f.+pl.]

According to the LLAR, the possible morphosyntactic constructions of the feminine plural DP are therefore the following (Rasom 2003):

1. determinantØ + adjectiveØ + noun-es
2. determinantØ + noun-es + adjective-es
3. determinantØ + nounØ + adjective-es

After presenting the different morphosyntactic combinations allowed in the LLAR, in the following section, I will include this phenomenon in the wider area of the studies on the Romance DP.

3. The dual source of adjectives in the Romance DP and their different semantic interpretations

Cinque’s work is based on two main generalisations, which will be analysed and discussed in order to demonstrate my hypothesis according to which the morphosyntax of the LLAR disambiguates the different interpretations of postnominal adjectives. The first generalisation deals with the possible readings of adjectives according to their prenominal or postnominal position: I will discuss this point in paragraph 3.1. The second generalisation corresponds to Cinque’s theory on the different syntactic sources of adjectives, depending on their position and on their reading. The latter will be discussed in paragraph 3.2.

3.1. The different order of adjectives in the Romance languages and their interpretation

As we can see in table 2, in Romance languages, whilst in the prenominal position (left column) the adjective has only one interpretation (individual, non-restrictive...) in the postnominal position (right column) it can have two interpretations: the same one of

prenominal adjectives or another one corresponding to individual level, non restrictive, relative (etc.) reading:

ADJECTIVAL INTERPRETATION		
INDIVIDUAL-LEVEL	N	STAGE-LEVEL OR INDIVIDUAL-LEVEL
NON-RESTRICTIVE	N	RESTRICTIVE OR NON-RESTRICTIVE
MODAL READING	N	MODAL READING OR IMPLICIT RELATIVE READING
NONINTERSECTIVE	N	INTERSECTIVE OR NONINTERSECTIVE
ABSOLUTE	N	RELATIVE OR ABSOLUTE
ABSOLUTE READING WITH SUPERLAT	N	RELATIVE OR ABSOLUTE READING WITH SUPERLATIVES
SPECIFICITY-INDUCING	N	SPECIFICITY-INDUCING OR NON-SPECIFICITY-INDUCING
EVALUATIVE	N	EVALUATIVE OR PROPOSITIONAL

Table 2 - The possible interpretations of prenominal and postnominal adjectives in Romance languages

We will now focus on some data of the different interpretations we dealt with in table 2 using Italian as an example of a Romance language. As for the interpretation of the oppositions, individual level/stage level (ex. 6-7), non-restrictive/restrictive (ex. 8-9), absolute/relative (ex. 10-11), we can see that if the adjectives ‘invisibili’, ‘noiose’, ‘alte e ripide’ precede the noun, (a.) is the only possible interpretation, whereas if the same adjective is in postnominal position, the interpretation can be the same of the prenominal adjective or the one in (b.). Here are the examples as listed in above.

Individual level / stage level

(6) Le invisibili stelle di Andromeda sono molto distanti.

The invisible stars of Andromeda are very far.

(a.) ‘Andromeda’s stars are all invisible and very far’.

(7) Le stelle invisibili di Andromeda sono molto distanti.

The stars invisible of Andromeda are all far.

(a.) ‘Andromeda’s stars are all invisible and very far’.

(b.) ‘there are some stars of Andromeda’s which are invisible and these are very far’.

Non-restrictive / restrictive

(8) Le noiose lezioni di Ferrari se le ricordano tutti.

The boring lessons of Ferrari refl – obj. pron. remember all.

(a.) ‘Ferrari’s lessons were all boring and all remember them’.

(9) Le lezioni noiose di Ferrari se le ricordano tutti.

The lessons boring of Ferrari refl – obj. pron. remember all.

(a.) ‘Ferrari’s lessons were all boring and all remember them’.

(b.) ‘all remember those lessons of Ferrari’s which were boring (but not all were so)’.

Absolute / relative

(10) Volevano scalare solo le alte e ripide montagne dell’India.

They wanted to climb only the tall and steep mountains of India.

(a.) ‘the mountains of India are all tall and steep and they wanted to climb them all’.

(11) Volevano scalare solo le montagne alte e ripide dell’India.

They wanted to climb only the mountains tall and steep of India.

(a.) ‘the mountains of India are all tall and steep and they wanted to climb them all’.

(b.) ‘they wanted to climb only those mountains of India which are tall and steep’.

These examples show that the interpretation of postnominal adjectives is ambiguous in Italian, as is in other Romance languages.

3.2. The dual source of adjectives: direct vs. indirect modification

The second main point discussed by Cinque in his work is that adnominal adjectives have two different sources corresponding exactly to the different adjectival readings, as was seen in paragraph 3.1. According to Cinque, there are two different possible positions occupied by the adjective in the syntactic tree. On the one hand, there are adjectives in **direct modification** source, that are APs (Adjective Phrases) generated in the Spec(ifier) of dedicated functional heads in the extended projection of the NP: these adjectives correspond to the readings individual level, non-restrictive, absolute (etc.). And on the other, there are adjectives in **indirect modification** source: these are reduced relative clauses (**Red**)RC and have stage level, restrictive, relative interpretation, as shown in the following syntactic tree:

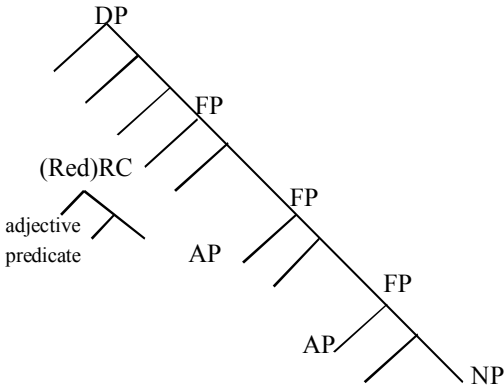


Figure 1 - The syntactic projection of the DP, according to Cinque’s theory

This tree shows how the adjective in (Red)RC position derives directly from the adjective predicate of a complete relative clause from which it inherits its semantic interpretation, whereas the APs at the bottom, which occupy a different position, also assume another semantic value.

4. The LLAR between morphosyntax and semantics: an evidence for Cinque's theory

In this paragraph, I will show how the Ladin varieties with LLAR differentiate, through morphology, the two sources of adjectives in postnominal position, i.e. in the position where the Romance adjective has an ambiguous interpretation.

As shown in section 2, the three possible constructions of the LLAR are the following:

- determinantØ + adjectiveØ + noun-es
- determinantØ + noun-es + adjective-es
- determinantØ + nounØ + adjective-es

The same examples of the different semantic interpretations of the adjective in the pronominal and postnominal positions will be presented (cf. 3.1), with the addition of the corresponding Ladin translations. As we will see, the different interpretations mirror perfectly in the morphosyntax of the LLAR.

As for the dichotomy individual level / stage level, in (12), (12a), where the order of the constituents is adjective plus noun, and in (13) and (13a), where the order is noun plus adjective, the interpretation is 'Andromeda's stars are all invisible and very far'. In (13b) the order remains noun plus adjective, yet the different morphology of the noun disambiguates the two interpretations in Ladin: in this case, the interpretation is that 'there are some stars of Andromeda's which are invisible and these are very far'.

With the order adjective plus noun in Italian and Ladin the possible morphological combinations are therefore the following:

ITALIAN = individual level

(12.) ITA. Le invisibili[+g+n] stelle[+g+n] di Andromeda sono molto distanti.

LADIN = individual level

(12a.) LAD. La steiles de Andromeda les é n muie dalènc.

The invisible [+g-n] stars [+g+n] of Andromeda cl are very far.

On the contrary, with the order noun plus adjective, the possible morphosyntactic combinations are the following:

ITALIAN = individual level, stage level

(13.) ITA. Le stelle[+g+n] invisibili[+g+n] di Andromeda sono molto distanti.

LADIN = individual level

(13a.) LAD. La steiles invisiboles de Andromeda les é n muie dalènc.

The stars[+g+n] invisibles[+g+n] of Andromeda cl are very far.

Ladin has this third morphological possibility:

LADIN = stage level

(13b.) LAD. La steila invisiboles de Andromeda les é n muie dalènc.

The star[+g-n] invisibles[+g+n] of Andromeda cl are very far

As for the dichotomy non-restrictive / restrictive, in (14), (14a), with the order adjective plus noun, and in (15), (15a), with the order noun plus adjective, the interpretation is that 'Ferrari's lessons were all boring and all remember them'. In Ladin there is another morphological possibility in (15b) which disambiguates the ambiguous interpretation of the postnominal adjective: 'everyone remembers those lessons of Ferrari's which were boring (but not all were so)'

With the order adjective plus noun, the possible morphological combinations are therefore the following:

ITALIAN = non restrictive

ITA. Le noiose[+g+n] lezioni[+g+n] di Ferrari se le ricordano tutti.

LADIN = non restrictive

(14a.) LAD. La stufousa ores de Ferrai i se les recorda duc.

The boring[+g-n] lessons[+g+n] of Ferrari cl pr. obj.pron remember all.

With the order noun plus adjective, on the contrary the possible morphosyntactic combinations are the following:

ITALIAN = non restrictive, restrictive

ITA. Le lezioni[+g+n] noiose[+g+n] di Ferrari se le ricordano tutti.

LADIN = non restrictive

(15a.) LAD.

La ores stufouses de Ferrai i se les recorda duc.

The lessons[+g+n] boring[+g+n] of Ferrari cl pr. obj.pron remember all.

LADIN = restrictive

(15b.) LAD.

La ora stufouses de Ferrai i se les recorda duc.

The lesson [+g-n] boring[+g+n] of Ferrari cl pr. obj.pron remember all.

As a last example, I will take into account the dichotomy absolute / relative reading: in (16), (16a), with prenominal adjective, and in (17), (17a), with postnominal adjective, the interpretation is that ‘the mountains of India are all tall and steep and they wanted to climb them all’. In (17b) again the order is noun plus adjective but the different morphological solution on the noun disambiguates the reading of the postnominal adjective: in this case therefore the interpretation is that ‘they wanted to climb only those mountains of India which are tall and steep’.

With the order adjective plus noun, the possible morphological combinations are the following:

ITALIAN = absolute

ITA. Volevano scalare solo le alte[+g+n] e ripide[+g+n] montagne[+g+n] dell’India.

LADIN = absolute

(16a.) LAD.

I volea demò rampeèr la auta e erta
crepes de l’India

*They wanted only to climb the tall[+g-n] and steep[+g-n]
mountains[+g+n] of India*

With the order noun plus adjective, on the contrary the possible morphosyntactic combinations are the following:

ITALIAN = absolute, relative

ITA. Volevano scalare solo le montagne[+g+n] alte[+g+n] e ripide [+g+n] dell’India.

LADIN = absolute

(17a.) LAD.

I	volea	demò	rampeèr la	crepes	autes
e	ertes		de l'India.		
<i>They</i>	<i>wanted</i>	<i>only</i>	<i>to climb the</i>	<i>mountains</i> [+g+n]	<i>tall</i> [+g-n]
<i>and</i>	<i>steep</i> [+g-n]		<i>of India</i>		

With the further Ladin solution:

LADIN = relative

(17b.) LAD.

I	volea	demò	rampeèr la	crepa	autes	e
ertes	de l'India.					
<i>They</i>	<i>wanted</i>	<i>only</i>	<i>to climb the</i>	<i>mountains</i> [+g-n]	<i>tall</i> [+g-n]	<i>and</i>
	<i>steep</i> [+g-n]		<i>of India.</i>			

On the bases of these examples, we can conclude that in the Romance languages with no lazy agreement, only syntax contributes to the semantic interpretation of adjectives, thus leaving the postnominal readings ambiguous. However, in the Ladin varieties with LLAR, morphology mediates in the disambiguation of the different postnominal readings.

5. General conclusions

Starting from the data presented in this work, the morphosyntactic study of the LLAR shows different points of interest: it gives further evidence to Cinque's theory on the interpretation of adjectives in prenominal and postnominal position; it shows how morphology mediated through syntactic movement can disambiguate potential ambiguous structures; it allows a more detailed split of the functional heads inside the DP; it invites to a comparative study of the data of other Romance varieties in order to find further generalisations about the morphosyntactic combinations inside the DP.

It should be mentioned that this work is only a part of a more thorough and wider analysis of the phenomenon of lazy agreement. In the attempt to describe a phenomenon of a minority language, the paper's main aim was to demonstrate the importance of these varieties in the survey of grammar and its internal morphosyntactic and semantic dynamics.

References

- CINQUE, Guglielmo (2005). *The dual source of adjectives and phrasal movement in the romance DP*. Ms., University of Venice, December 2005.
- HAIMAN, John & BENINCÀ, Paola (1992). *The Raetho-Romance Languages*. London, New York, N.Y.: Routledge, VIII.
- RASOM, Sabrina (2003). Una particolarità del plurale femminile nel ladino dolomitico. Considerazioni morfosintattiche, *Atti del Convegno internazionale di studi*. Sappada/Plodn (BL), 2-6 luglio 2003: 239-246.
- RASOM, Sabrina (2005). Il plurale femminile nel ladino dolomitico tra morfologia e sintassi, *Quaderni di Lavoro Asis (on line)*, 5: 20-35.

LES CRITÈRES DE GRAMMATICALISATION POUR LES TEMPS COMPOSÉS EN FRANÇAIS ET EN ALLEMAND

Aude REBOTIER

aude@rebotier.com

Xerox Research Centre Europe

6 chemin de Maupertuis

38240 Meylan

Abstract : This article examines two states of grammaticalization of compound verbal forms: Phase and Tense. Criteria are proposed for determining whether a particular form is to be considered as conveying Phase or Tense. These criteria are then applied to French and German structures. Our results show that no uniform criteria can be set up, and that the grammaticalization process depends on the meaning of the components.

On admet en général que la composition est un moyen de former de nouveaux Temps : certaines associations de deux formes verbales, l'une finie et l'autre non finie (infinitif ou participe), soumises à un processus de grammaticalisation, pourraient accéder au statut de Temps ; la forme finie cesse alors d'être un lexème pour devenir un auxiliaire. Je commencerai par préciser la notion de grammaticalisation, et quels critères on peut envisager pour la mesurer, puis j'exposerai les résultats des tests.

1. Le processus de grammaticalisation

1.1. L'exemple des parfaits

Dahl (1995) envisage pour toutes les langues un noyau qui contient les formes verbales tout à fait grammaticalisées, et une périphérie où se trouvent les formes en voie de grammaticalisation, le plus souvent composées, qui tendent à entrer dans le noyau¹. Les Parfaits (c'est-à-dire les formes composées du verbe *être/avoir* + participe passé) tendent ainsi à devenir des Temps passés² ou des perfectifs, tandis que les Futurs proviennent de diverses expressions lexicales exprimant la modalité (volonté / obligation comme en anglais : *will/shall*), le mouvement (*aller* en français ou *to be going to* en anglais) ou encore le changement (*werden* en allemand). Ces expressions passeraient par un stade intermédiaire en périphérie avant de devenir, éventuellement, de véritables Temps futurs.

On peut distinguer des étapes dans le processus de grammaticalisation. Les Parfaits en offrent un bon exemple.

Avec un verbe *avoir* ou *être*, on peut utiliser les participes passés de manière attributive : attribut du sujet (1, 2) ou de l'objet (3, 4).

1. Die Fesseln des Pferdes **sind bandagiert**.
2. Les paturons du cheval **sont bandés**.

¹ Dahl (1985) propose une approche comparative entre plus de 60 langues sur les catégories de Temps et d'Aspect.

² Lorsque le mot est ambigu en français, la majuscule signale les formes grammaticales : Temps (Tense, Tempus), Futur, Présent, par opposition à temps (Time, Zeit), futur (= avenir), présent (= moment présent). Futur, FS (Futur simple), FP (Futur proche), Parfait/Perfekt, PC (passé composé), PS (passé simple) désignent des formes verbales indépendamment de leur sens : elles peuvent relever de la catégorie Temps, de la catégorie Phase/Aspect, ou ne pas être grammaticalisées.

3. Das Pferd **hat** die Fesseln **bandagiert**. (cf. Vater, 1994 : 60)

4. Le cheval **a** les paturons **bandés**.

Dans ce cas, en français, le participe est substituable par un adjectif :

5. Le cheval **a** les paturons sales.

Ces constructions se distinguent de véritables Parfaits comme dans les énoncés suivants :

6. Der Knecht **hat** die Fesseln des Pferdes **bandagiert**.

7. Le palefrenier **a bandé** les paturons du cheval.

par des phénomènes syntaxiques (en français, la position du participe, l'accord avec le sujet ou l'objet) et par des phénomènes sémantiques : avec un Parfait, il y a toujours un procès passé (ou antérieur à un moment), réalisé par le sujet grammatical, ici le palefrenier. Ces conditions ne sont pas nécessairement remplies lorsque le participe n'appartient pas à une forme de Parfait (par exemple, *une région éloignée* est une région lointaine, mais pas une région qu'on a (ou qui s'est) éloignée).

Lorsqu'on a affaire à un véritable Parfait, il faut encore distinguer s'il relève du Temps ou de la Phase. Tout énoncé parle d'un intervalle temporel qu'on peut appeler son *moment thématique* (concept inspiré du *topic time* de Klein, 1994 : 3, 36-58), et réalise une prédication de ce moment. Je pose qu'un Temps situe le moment thématique dans le temps³. Le moment thématique peut être explicité ou précisé par un complément temporel. Avec un Parfait, le complément peut dater soit le procès, soit l'état qui suit le procès, et qu'on peut appeler sa phase résultante :

8. Pierre **avait quitté** son bureau à cinq heures. (d'après He, 1997 : 130ss)⁴

9. À cinq heures, Pierre **avait déjà quitté** son bureau.

Dans ces deux exemples, il est question du procès *quitter son bureau* ; le résultat est l'absence de Pierre. Le complément *à cinq heures* peut situer le départ de Pierre (8) ou son absence, l'état résultant (9). Dans le premier cas, on peut considérer le Plqpf comme un Temps : la phrase situe le moment thématique dans le passé, et le procès *quitter son bureau* directement par rapport à ce moment ; dans l'énoncé 9, le Temps de la phrase est l'Impft, et la forme de Parfait (*avoir* + participe) est une expression de Phase, l'accompli : elle indique que ce qui est mis en relation par la prédication avec le moment thématique n'est pas le procès exprimé par l'infinitif, mais son résultat⁵. On peut faire l'hypothèse que l'expression de la Phase est une première étape de grammaticalisation avant le Temps (cf. Bybee, Perkins & Pagliuca, 1994 : 51-105) : comme dans les énoncés 1 à 5, c'est le Temps d'*avoir/être* qui fixe le moment thématique ; dans l'exemple 8 en revanche, c'est l'ensemble auxiliaire + participe.

Les Parfaits existent dans de nombreuses langues, à différents stades de grammaticalisation. En anglais, il est difficile de dater le procès avec un *present perfect* : la forme n'exprimerait

³ Le Présent, par exemple, indique généralement que l'énoncé parle du moment d'énonciation, et que le procès est vrai pour ce moment, quelle que soit par ailleurs l'étendue du procès.

⁴ Un complément rhématique comme *à cinq heures* dans l'exemple 8 ne définit pas le moment thématique, mais le moment du procès ; cependant, celui-ci est ici inclus dans le moment thématique, qu'on peut expliciter par ailleurs : « La veille, il avait quitté son bureau à cinq heures. »

⁵ Si l'on interprète le point R de Reichenbach (1947) comme un moment thématique, ce modèle est compatible avec une conception répandue du Temps comme relation entre S et R, tandis que l'aspect serait une relation temporelle entre E et R (Comrie 1976, Klein 1994, Gosselin 1996). Je préfère cependant réserver le terme d'aspect à des phénomènes (comme l'opposition Impft/PS) qui sont en correspondance directe avec l'aspect lexical, et utiliser celui de Phase pour le déplacement du moment thématique avant ou après le procès (cf. Rebotier, 2005 : 16-23 et 108-226). L'accompli est aussi appelé résultatif ou antérieur (Bybee, Perkins & Pagliuca, 1994).

que la Phase. En russe ou en polonais au contraire, l'évolution serait achevée : l'ancien Parfait est devenu l'unique forme de passé (Veyrenc, 1970 : 76-82). Entre les deux, les Parfaits seraient en français comme en allemand des formes intermédiaires, capables d'exprimer la Phase ou le Temps (He, 1997, Vater, 1994). L'approche contrastive est ici particulièrement intéressante, car elle permet de montrer différents stades de grammaticalisation pour des formes similaires ; elle permet aussi d'établir des critères moins spécifiques aux formes que l'on trouve dans une langue donnée.

Pour la phase, le pendant de l'accompli est le prospectif : il réalise une prédication sur la période qui précède le procès et non sur le procès lui-même. Mais la phase préparatoire est moins facile à caractériser que l'état qui résulte d'un procès : on voit bien quel est l'état résultant du procès *quitter le bureau*, il est plus difficile de dire où commence la préparation à ce procès.

1.2. Le projet d'étude

1.2.1. Les formes étudiées

En français et en allemand, plusieurs formes se composent d'un verbe fini et d'une forme verbale non finie (infinitif ou participe) :

- dans les deux langues, les formes de Parfaits (Passé Composé : PC, et Perfekt) ;
- en français, une autre forme pour le passé : venir de + infinitif ;
- en allemand, la forme appelée traditionnellement Futur est une forme composée (*werden*, verbe qui signifie *devenir*) ;
- en français, le Futur simple est une ancienne forme composée de l'infinitif et de l'auxiliaire *avoir*, et le Futur Proche (FP) est formé sur le verbe *aller* ;
- dans les deux langues, certains verbes de modalité peuvent entrer dans des compositions qui font office de futurs : en allemand, *wollen* (vouloir) et *sollen* (devoir) ; en français, *devoir*, notamment à l'Impf pour faire au cours d'un récit des prédictions dont on sait qu'elles se sont réalisées par la suite⁶.
- être sur le point de + infinitif est une forme qui n'est jamais comptée parmi les Temps (qui aurait donc au mieux atteint le stade de la Phase) et qui offre un point de comparaison.

1.2.2. Les critères de grammaticalisation

Il s'agit de déterminer à quel degré ces formes composées sont grammaticalisées : sont-elles de véritables Temps, expriment-elles la Phase, ou bien ne sont-elles que la composition d'éléments qui gardent chacun tout leur sens ?

- Un premier test (2.1) consiste à introduire un complément de temps : peut-il dater le procès, ou au contraire sa phase préparatoire ou son résultat ?
- Un deuxième critère (2.2) est valable pour toute grammaticalisation : le sens de la forme est-il déductible du sens de ses parties, autrement dit, le sens de l'auxiliaire est-il encore perceptible ?
- Les possibilités de conjuguer l'auxiliaire d'une forme soupçonnée d'être un Temps devraient être réduites (2.3)⁷. Les expressions lexicales en revanche, ainsi que l'expression de la Phase, peuvent s'appliquer à tout le système verbal.

⁶ Prophétie rétrospective, comme dans l'exemple suivant :

En 1831, le jeune état couronnait son premier roi, Leopold de Saxe-Cobourg. Il **devait consacrer** la plupart de sa vie à la consolidation du nouveau pays.

(http://www.belinfo.be/fr/belgique/belgique_histoire.htm, consulté le 03/04/2005)

⁷ Dans la tradition guillaumienne, reprise par exemple par Wunderli (1976 : 10), la distinction des Temps est l'apport propre de l'indicatif.

- La Phase impose au procès des limitations aspectuelles (2.4).

On pourrait ajouter un critère de fréquence : on exige parfois qu'une forme verbale soit obligatoire dans certaines circonstances pour pouvoir être qualifiée de Temps. Je ne développerai pas ce dernier critère, car mon corpus d'étude n'est pas assez vaste.

En tout, j'ai soumis neuf formes composées à six tests : je ne donnerai ici qu'un aperçu des résultats.

2. Résultats des tests

2.1. Datation

L'introduction d'un complément de temps permet de distinguer trois types de formes composées.

a. Certaines formes n'acceptent que la datation du procès : Futur simple et *werden* + infinitif.

10. À cinq heures, Pierre **quittera** son bureau.
11. Pierre **quittera** son bureau à cinq heures.
12. Peter **wird um fünf** sein Büro **verlassen**.
13. Um fünf **wird** Peter sein Büro **verlassen**.

Dans ces quatre énoncés, c'est bien le procès lui-même (le départ de Pierre) qui doit avoir lieu à cinq heures, et non une phase préparatoire.

On peut ajouter les verbes de modalité : lorsqu'ils sont employés comme des quasi-futurs, on considère la modalité comme une condition suffisante pour que le procès se réalise effectivement :

14. Wir **wollen** es bei dir **einstellen**. (B. Brecht, *Der gute Mensch von Sezuan*)
15. Nous allons le remiser chez toi.

Au contraire d'une traduction littérale en français (*nous voulons le remiser chez toi*), l'énoncé 14 affirme bien (dans son contexte) que l'action n'est pas seulement souhaitée, mais prévisible pour l'avenir, ce qui doit se rendre par un Futur en français, ici un FP (15). Cet emploi est limité à certaines conditions (verbe au Présent, première personne) ; en particulier, il n'est pas compatible avec une datation de la modalité :

16. Heute **wollen** wir es bei dir **einstellen**.
17. Aujourd'hui, nous **voulons** le remiser chez toi.

Dans cet énoncé modifié, *heute* peut dater le procès à l'infinitif (il s'agit de *le remiser aujourd'hui*) ; s'il date la modalité (*nous le voulons aujourd'hui*), elle est mise en valeur, et la réalisation du procès n'apparaît plus comme sa conséquence naturelle : ce n'est plus une phase préparatoire mais une simple modalité comme en français. Enfin, dans le cas de la prophétie rétrospective, il n'y a pas de phase préparatoire : le principe d'une prophétie consiste à affirmer la réalisation d'un événement dans l'avenir sans le mettre en relation avec le présent (ou le passé).

b. A l'inverse, *être sur le point de* n'accepte que la datation de la phase préparatoire :

18. A cinq heures, Pierre **était sur le point de quitter** son bureau.
19. Pierre **était sur le point de quitter** son bureau à cinq heures.

Dans les deux énoncés, *cinq heures* n'est pas l'heure où Pierre a quitté son bureau, mais où il se prépare à le faire.

Venir de est parfois compatible avec la datation du procès, mais c'est rare, et en général entre virgules, comme un commentaire mal intégré syntaxiquement⁸ :

⁸ En relevant les formes *venir de* + infinitif dans 18 journaux francophones (18-30 juin 2006), j'ai trouvé sur 484 occurrences 17 compléments de temps ; 15 sont soit séparés du verbe par une virgule

20. le Sénat américain **vient de voter**, à la fin mai, un texte proposant la légalisation de sept millions d'immigrés. (Le Figaro, 30/06/06)

c. Enfin, deux formes acceptent les deux datations : les Parfaits (on l'a vu sous 1.1), et le FP qui, symétriquement, a des emplois de prospectif (21) et de Futur (22) (cf. Vet, 1985) :

21. Laforge l'arrêta au moment où il **allait passer** la porte. (Daniel Pennac)

22. j'ai peur du moment où je **vais** vraiment **me retrouver** toute seule avec un appart et tout a gerer... (Internet, <http://forum.doctissimo.fr/psychologie/reves-cauchemars/sujet-148923.htm>, consulté le 5 avril 2005)

Dans l'exemple 21, *le moment où il allait passer la porte* est le moment qui précède le procès (le passage). En revanche, dans 22, *le moment où je vais me retrouver seule* est bien le moment où le procès (se retrouver seule) se réalise ; dans ce cas, il est possible de substituer un FS (ou un conditionnel s'il s'agit de passé) au FP ; ce n'est pas possible dans les emplois de Phase⁹ :

23. ?? Laforge l'arrêta au moment où il **passerait** la porte.

24. J'appréhende le moment où je **me retrouverai** seule.

Cependant, il existe des états intermédiaires entre Phase et Temps. J'en donnerai un seul exemple, celui des Parfaits de bilan.

Dans certains énoncés, un complément en *depuis* ne donne pas la durée du procès résultant, mais fixe un cadre pour le procès, un intervalle qui s'étend jusqu'au présent et au cours duquel le procès a lieu :

25. Depuis le début du siècle, une discipline scientifique **a été développée** pour répondre à ce besoin.

Par cet énoncé, on apprend seulement que la discipline n'était pas développée au début du siècle et qu'elle l'est aujourd'hui, on ignore quand exactement elle s'est développée entre les deux : il s'agit d'un bilan sur cette période. C'est un emploi de Temps, puisque c'est le procès qui est situé dans le temps, mais en relation avec un état résultant actuel.

Ce premier test montre donc que certaines formes n'acceptent que la datation de la période qui suit ou qui précède le procès, condition pour parler de Phase ; que d'autres, les plus grammaticalisées, n'acceptent que la datation du procès, condition pour parler d'un Temps ; et qu'un troisième groupe accepte les deux datations, à cheval entre les deux catégories.

2.2. Sens de l'auxiliaire

La grammaticalisation s'accompagne de la perte du sens plein de l'auxiliaire. Dans le FP par exemple, *aller* perd son sens de déplacement physique. Le phénomène est bien visible car les deux emplois existent :

26. Où vas-tu ? – Je **vais prendre** le train.

27. Je **vais** enfin **savoir** la vérité.

Dans l'exemple 26, on peut distinguer deux procès successifs, celui d'aller (à un endroit) et celui de prendre le train. En revanche, dans l'exemple 27, un FP, le locuteur n'envisage pas de se déplacer pour savoir la vérité.

Un indice de la perte du sens plein est l'impossibilité d'appliquer la forme composée à son auxiliaire : ce n'est possible que si l'auxiliaire a perdu son sens plein, sinon il y a

(comme l'exemple 20), soit dans des relatives. Je n'ai trouvé que deux exemples de dates sans virgule en proposition indépendante ou principale.

⁹ La substitution par un Temps simple est également utilisée pour distinguer les emplois de Phase et de Temps du Parfait (Ballweg, 1988 ; Latzel, 1997). Elle est cependant à prendre avec précautions, car deux Temps ne sont pas nécessairement interchangeables dans tous les contextes.

redondance. Ainsi, « Je vais aller » ne peut être qu'un FP : l'auxiliaire *aller* n'a pas le sens plein du verbe *aller* mis ici au FP.

Les résultats de ce test sont à peu près cohérents avec la possibilité de dater le procès, sauf pour les verbes de modalité : il semble difficile de lire « ich will wollen » ou « il devait devoir » dans le sens d'une annonce concernant l'avenir (ou le futur du passé).

On peut remarquer aussi que le test est négatif pour *venir de* : il est difficile de dire « il vient de venir », alors même que le sens plein de *venir* (déplacement physique) est perdu, ce qui confirme les réticences qui semblaient apparaître pour dater le procès.

Ce test ne dit pas si un sens figuré est encore perceptible : *aller* peut exprimer non un déplacement physique, mais le fait que la situation progresse vers le procès à venir : c'est le propre de la Phase. Ces sens affaiblis sont beaucoup plus difficiles à estimer.

2.3. Défectivité

D'après Wunderli (1976), les formes composées romanes qui deviennent des Futurs limitent la conjugaison de leur auxiliaire au Présent et à l'Impft. En tout cas, on peut attendre d'un Temps l'incompatibilité avec les formes non finies : il n'y a pas dans nos langues d'infinitif Futur, ni d'infinitif Impft ou Prétérit.

Les Futurs du français sont effectivement limités au Présent et à l'Impft (le conditionnel représentant un FS dont l'auxiliaire fusionné est à l'Impft). Lorsque l'auxiliaire *aller* du FP est au PS ou à l'infinitif, il reprend son sens plein et implique un déplacement physique (*il alla s'exprimer*).

Le Futur allemand accepte le subjonctif 1 et 2, contrairement au Prétérit : mais il faut avoir à l'esprit que le subjonctif allemand n'a pas la même fonction que le subjonctif français, et qu'il est un calque obligatoire des formes de l'indicatif dans certains contextes. En revanche, le prétérit de la formule (*wurde gehen*), forme historique, n'existe plus, et l'infinitif de *werden* + infinitif, bien que facile à construire, est une forme impossible.

Un deuxième groupe, qui réunit les Parfaits et *être sur le point de*, ne connaît pas de défectivité et traverse tout le système verbal. Enfin, certaines formes ont une défectivité limitée : en français, *devoir* + infinitif est précisément utilisé comme substitut du Futur lorsque celui-ci est impossible, par exemple dans une subordonnée en *si*, ou encore à l'infinitif :

28. Je ne signerai pas le contrat s'il **doit contenir** une pareille clause. (Imbs, 1960 : 58)

En revanche le PS ou le PC empêche tout emploi prospectif : « il dut partir » signifie que le sujet est vraiment parti, et non qu'il allait partir plus tard. La défectivité est donc limitée à ces deux Temps.

On constate que les trois tests ne coïncident pas exactement ; en particulier, la double datation, qui concerne les Parfaits comme le FP, n'implique rien quant à la défectivité.

2.4. Aspectualité

Un autre test pour distinguer Phase et Temps consiste à examiner si la forme peut exprimer un procès en continuité avec le présent. En effet, lorsqu'on utilise l'accompli, on distingue le moment du procès et sa phase résultante, qui commence ensuite. Ceci implique que le procès soit terminé, on peut d'ailleurs dater sa fin, qui est le début de l'état résultant (« il est parti depuis une heure »). De la même façon, au prospectif, le procès est précédé d'une phase préparatoire pendant laquelle il n'a pas encore commencé. Il n'est donc pas dans la continuité du présent. On constate que la continuité est possible pour les Futurs et les Parfaits :

29. C'est ainsi et ce **sera** toujours ainsi.

30. Es ist so und **wird** immer so **sein**.

31. Cela **a** toujours **été** ainsi.

En revanche elle est impossible pour les formes qui ont déjà répondu négativement aux autres tests :

32. ?? C'est ainsi et c'**est sur le point d'être** toujours ainsi.

33. ?? Cela **vient d'être** toujours ainsi

Et la formule est douteuse pour le FP :

34. ? C'est ainsi et cela **va être** toujours ainsi.

Une autre conséquence de la Phase est qu'on ne peut pas voir le procès de l'intérieur, puisqu'on le perçoit indirectement, depuis sa phase préparatoire ou résultante¹⁰. L'incidence sur le modèle de l'Impft est donc impossible :

35. Quand je suis entré, il **était** à son bureau.

Cet énoncé ne dit pas quand le sujet s'est mis à son bureau, ni s'il y est encore.

Le Futur simple et le Futur allemand acceptent bien ces emplois, le FP, plus difficilement :

36. Quand j'entrerai, il **sera** à son bureau.

37. ? Quand je vais entrer, il **va être** à son bureau.

Et *venir de* et *être sur le point de* ne l'acceptent pas du tout :

38. * Il **vient d'être** à son bureau quand je suis entré.

Les Parfaits se comportent différemment : l'incidence est possible en allemand, alors que le même emploi est impossible en français :

39. Als ich ins Zimmer gekommen bin **hat** er am Schreibtisch **gesessen**.

40. ? Quand je suis entré, il **a été** à son bureau.

On a dans l'exemple 40 au mieux deux procès successifs, et non une incidence.

En allemand, ces emplois sont le signe d'une grammaticalisation en cours : le Parfait peut jouer le rôle du Prétérit dans certains cas. En français en revanche, cette impossibilité n'indique pas une grammaticalisation moins avancée, puisque les Temps du passé sont porteurs d'aspect en français ; dans un schéma d'incidence, le procès incident doit être au PS ou au PC, le procès auquel il est incident doit être à l'Impft. Le PC, en se grammaticalisant, prend logiquement le rôle du PS dans son opposition à l'Impft.

3. Conclusion

	datation	double auxiliaire	défectivité	continuité	incidence
<i>être sur le point de</i> + inf.	phase	non	non	non	non
<i>venir de</i> + inf.	(phase)	non	limitée	non	non
parfait français	procès/phase	oui	non	non	non
parfait allemand	procès/phase	oui	non	non	oui
Futur Proche	procès/phase	oui	oui	(oui)	(non)
verbes modaux + inf.	procès	non	(non)	oui	oui
<i>werden</i> + inf.	procès	oui	oui	oui	oui
Futur simple	procès	oui	oui	oui	oui

Tableau 1 - application de critères de grammaticalisation aux formes verbales composées

¹⁰ Ces limitations aspectuelles ont naturellement des conséquences sur l'aspect lexical, que je ne développerai pas ici.

En conclusion, la grammaticalisation n'est pas un processus uniforme. Les critères varient d'une forme composée à l'autre. Ainsi, les verbes de modalité ne passent pas par l'expression de la Phase : d'après les tests, soit ils se comportent comme des Futurs, soit ils ont leur sens plein. On constate aussi une dissymétrie entre le passé et l'avenir : la défektivité par exemple ne semble pas être un critère pertinent pour les formes du passé. D'autre part le prospectif ne semble pas être un véritable correspondant de l'accompli : le FP se comporte très différemment des Parfaits.

Le tableau montre cependant que la forme allemande *werden* + infinitif, très contestée comme Temps (entre autres à cause de sa forme composée), a toutes les caractéristiques formelles d'un Temps futur (cf. Rebotier, 2006). Enfin on ne constate pas de véritable différence entre les Parfaits allemand et français, alors qu'on attendait une grammaticalisation plus avancée en français à cause de son emploi généralisé à l'oral, par défaut du PS : les tests ne montrent certainement pas toutes les nuances de grammaticalisation¹¹.

Références

- BALLWEG, Joachim (1988). *Die Semantik der deutschen Tempusformen : eine indirekte Analyse im Rahmen einer temporal erweiterten Aussagelogik*. Düsseldorf : Schwann.
- BYBEE, Joan, PERKINS, Revere & PAGLIUCA, William (1994). *The Evolution of grammar*. Chicago, Londres : The University of Chicago Press.
- COMRIE, Bernard (1976). *Aspect. An Introduction to the study of Verbal Aspect and related Problems*, Cambridge : Cambridge University Press.
- DAHL, Östen (1985). *Tense and Aspect Systems*, Oxford : Blackwell.
- DAHL, Östen (1995). Das Tempussystem des Deutschen im typologischen Vergleich, *Deutsch - typologisch*. (= *Jahrbuch Institut für deutsche Sprache* 1995), LANG, Ewald & ZIFONUN, Gisela (Eds). Berlin/New York : de Gruyter, 359-368.
- GOSSELIN, Laurent (1996). *Sémantique de la temporalité en français*, Louvain-la-Neuve : Duculot.
- HE, Robert (1997). Vers une nouvelle vue d'ensemble des aspects en français, *Cahiers de grammaire*, 22 : 127-153.
- IMBS, Paul (1960). *L'emploi des temps verbaux en français moderne*, 4^e édition (1965) Paris : Klincksieck.
- KLEIN, Wolfgang (1994). *Time in language*, Londres/New York : Routledge.
- LATZEL, Sigbert (1997). Perfekt und Präteritum in einem deutschen Zeitungstext, *Nouveaux Cahiers d'Allemand*, 2 : 229-236.
- REBOTIER, Aude (2005). *Les Temps verbaux en allemand et en français et les phénomènes aspectuels. Comparaison de deux systèmes*, thèse de doctorat. Grenoble 3.
- REBOTIER, Aude (2006). Le Futur de l'allemand, *CORELA*, 4, 1 (<http://edel.univ-poitiers.fr/corela/document.php?id=1172>).
- REICHENBACH, Hans (1947). *Elements of Symbolic Logic*, New York/London: The Free Press/Collier-Macmillan.
- VATER, Heinz (1994). *Einführung in die Zeit-Linguistik*. Hürth-Efferen : Gabel Verlag.
- VET, Co (1985). Univers de discours et univers d'énonciation : les temps du passé et du futur, *Langue française*, 67 : 38-58.
- VEYRENC (1970), Charles Jacques. *Histoire de la langue russe*, Paris : PUF (= *QSJ* 1368).
- WUNDERLI, Peter (1976). *Tempus und Modus. Beiträge zur synchronischen und diachronischen Morphosyntax der romanischen Sprachen*. Tübingen : Narr.

¹¹ Cet article reprend pour l'essentiel la troisième partie de ma thèse : Rebotier (2005 : 228-329).

ÉTUDE TYPOLOGIQUE DE L'EXPRESSION DU DÉPLACEMENT EN ARABE TUNISIEN

Darine SAÏDI

saidi_dareen@yahoo.fr

UMR 5596 - CNRS & Université Lumière Lyon 2

Laboratoire Dynamique du Langage

Institut des Sciences de l'Homme

14, avenue Berthelot

69363 LYON CEDEX 07

Abstract : Following Talmy's work (1975, 1985, 2000) in the domain of Motion Events, languages fall into two types according to how they lexicalize directed motion events. Talmy's typology has led researchers to classify languages as either Satellite or verb-framed languages, although they do not fit all comfortably into this dichotomy. As a Semitic language, Tunisian Arabic is considered a verb-framed language. However, other possibilities do exist for the encoding of motion events in this language.

1. Présentation de la langue

1.1. *Situation linguistique actuelle en Tunisie*

Le passé linguistique de la Tunisie est caractérisé par des contacts linguistiques incessants plus ou moins intenses, puis par la colonisation française qui ont abouti à l'instauration d'un véritable bilinguisme arabo-français se greffant sur une situation de diglossie dialecto-littérale à clivage très net. L'époque post-coloniale est encore politiquement instable dans les choix fondamentaux, notamment dans le domaine linguistique. Tout cela ne peut aboutir qu'à une situation linguistique fluctuante mais nécessairement évolutive et dynamique.

1.2. *Langues et niveaux de langue en présence*

1.2.1. L'arabe littéraire

Il s'agit d'un niveau écrit, lu et secondairement parlé (dans l'enseignement et quelques rares discours improvisés). C'est la langue de la littérature en général, des journaux, des revues, de certaines émissions de radio et de télévision (notamment les informations et les reportages).

1.2.2. L'arabe dialectal

Il s'agit d'un niveau essentiellement parlé, qui représente la langue pratiquée par tout le peuple tunisien, (abstraction faite du parler berbère dont le champ est extrêmement réduit, à peine 1% de la population tunisienne). Il y a certes des variantes régionales au niveau phonologique et surtout lexical mais elles ne présentent aucun obstacle à l'intercompréhension. Ce niveau oral est cependant écrit et lu dans certaines émissions radiophoniques et même dans le domaine littéraire, notamment le théâtre et les dialogues dans certains romans. Le dialectal et le littéral diffèrent aux niveaux phonologique, morphologique et lexical. La traduction d'une même séquence donne deux modes de réalisation différents : le littéral s'en tient à un code fixe alors que le dialectal utilise différents mécanismes (lexicalisation, analogie, métaphorisation, etc.) pour la dénomination de réalités nouvelles. On observe une grande créativité dans le domaine du dialectal en raison de nombreux emprunts, de l'adaptation à de nouvelles réalités et de son caractère de

langue orale non normée. D'où l'intérêt de se pencher sur l'arabe dialectal tunisien, d'autant plus que très peu d'études sont consacrées à ce parler.

1.2.3. Le français

Le français n'est pas simplement la langue de l'ouverture et de la modernité, comme on le qualifie, mais c'est une langue véhiculaire et de prestige dans l'usage quotidien. La situation linguistique en Tunisie s'avère très complexe, se caractérisant d'une part par une diglossie arabe langue écrite-langue orale (les deux niveaux littéral et dialectal ayant chacun leur propre système), d'autre part par un bilinguisme (arabe – français). Tout au long de son histoire et jusqu'à présent l'arabe tunisien a évolué avec d'autres langues qui ne sont pas nécessairement de la même famille. Cette situation influe certainement sur la structure de la langue.

Nous nous intéresserons dans cet article à l'arabe tunisien dialectal que nous appellerons désormais arabe tunisien. En effet, nous avons déjà souligné le dynamisme de ce système et le fait qu'il a été extrêmement peu étudié. Au sein de ce système nous étudierons ici l'expression du déplacement et cela au travers de la typologie de l'expression du déplacement de Talmy.

2. Typologie du déplacement selon Talmy

L'étude de l'expression du déplacement (relation spatiale dynamique) examine la manière dont une langue encode linguistiquement ce domaine sémantique particulier. L'objectif principal de cette typologie est d'une part d'examiner les relations systématiques entre les différents éléments sémantiques présents dans un événement spatial (trajectoire, fond, manière, etc.) et, d'autre part, de classer les langues selon les outils morphosyntaxiques employés pour encoder la notion de trajectoire (Talmy, 1991). Il s'agit aussi d'identifier les préférences des langues quant à la lexicalisation des éléments sémantiques encodés dans le verbe principal d'une phrase (Talmy, 1985, 2000).

2.1. Types de lexicalisation

Talmy montre à quel point les langues diffèrent dans la façon d'encoder les différentes composantes sémantiques associées au déplacement (trajectoire, manière ou cause, figure et fond). La figure est définie comme l'entité qui se déplace et le fond comme l'entité de référence. Suite à un examen de la structure sémantique des verbes de déplacement dans des langues souvent très éloignées, Talmy dégage trois types de lexicalisation :

2.1.1. Type 1 : V [déplacement + trajectoire]

Lexicaliser dans la racine verbale le déplacement et la trajectoire. Talmy (2000 : 49) donne un exemple présentatif en espagnol, celui de deux verbes dits de trajectoire : *entrar* 'entrer' qui exprime le déplacement dirigé de l'extérieur vers l'intérieur d'un lieu et le verbe *salir* 'sortir' qui exprime le déplacement orienté de l'intérieur vers l'extérieur.

- 1a. La botella entró a la cueva (flotando)
'La bouteille est entrée dans la grotte en flottant.'
- 1b. La botella salió de la cueva (flotando)
'La bouteille est sortie de la grotte en flottant.'

2.1.2. Type 2 : V [déplacement + manière ou cause]

Ce type de lexicalisation se trouve par exemple en anglais, il consiste à lexicaliser dans la racine du verbe le déplacement et la manière et/ou la cause du déplacement, exemple le verbe *slide* 'glisser' en anglais peut encoder la manière d'un déplacement continu sans secousse le long d'une surface :

2. The rock slid down the hill. (Talmy, 2000 : 28)
 'Le rocher a glissé le long de la colline.' (Traduction des éditeurs)

2.1.3. Type 3 : V [déplacement + figure]

Encoder dans la racine du verbe le déplacement et l'entité qui se déplace ou qui est déplacée (appelée figure). C'est un type de lexicalisation rare dans les langues du monde dans la mesure où le verbe n'encode pas systématiquement déplacement + figure. Talmy donne l'exemple des verbes anglais (3) to rain 'pleuvoir' et to snow 'neiger' qui combinent dans leurs racines deux constituants sémantiques, la figure et le déplacement de celle-ci avec une orientation vers le bas.

Les langues sémitiques, et donc l'arabe tunisien, feraient partie, selon le classement typologique de Talmy, des langues de type 1, c'est-à-dire qui lexicalisent dans la racine verbale le **déplacement et la trajectoire**. Ce classement pose déjà problème car le classement des « langues sémitiques » en général, repose probablement sur la base de l'étude d'une seule de ces langues (l'hébreu). En effet, d'autres familles linguistiques se répartissent entre plusieurs stratégies. Quand bien même on parlerait plus précisément de l'arabe à l'intérieur des langues sémitiques, on ne sait pas de quel arabe on parle, le littéral ou le dialectal et si c'est le dialectal, celui de quel pays ou de quelle région...

Au premier abord, le tunisien semble bien être conforme au classement typologique de Talmy puisque, comme le montre l'exemple (4), le déplacement et la trajectoire sont lexicalisés dans le verbe. Le tunisien possède bien des verbes appartenant au type 1 (langues qui lexicalisent dans un seul verbe le déplacement et la trajectoire).

- | | | | |
|--------------------------------|------------------|----|-----------|
| 4. it-t ^o fol | χræ3 | mi | l-mædræ |
| DEF-enfant | PAST-sortir.3MSG | de | DEF-école |
| L'enfant est sorti de l'école. | | | |

Cependant, un second problème (après celui du classement en bloc des langues sémitiques) est qu'il existe bien, comme on le verra, d'autres stratégies de lexicalisation en arabe tunisien qui ne trouvent pas de place dans cette typologie ternaire de Talmy dans laquelle les différents types ne sont pas exclusifs.

2.2. Langues à cadre verbal versus langues à satellites

Nous venons de voir une première typologie qui s'intéresse à l'encodage des différents éléments sémantiques dans le verbe. Nous allons voir maintenant une deuxième typologie qui s'intéresse à l'encodage de la trajectoire. La typologie de l'expression du déplacement (*typology of motion event*) proposée par Talmy (1991, 2000) se base sur la distribution morphosyntaxique des éléments sémantiques employés par la langue pour décrire la trajectoire. On distingue :

- Les langues à cadre verbal qui encodent la trajectoire dans un verbe principal et la manière dans un syntagme associé (*verb-framed language*) ex : l'espagnol, les langues romanes et les langues sémitiques (selon Talmy) dont fait partie l'arabe tunisien.
5. L'enfant sort de l'école en courant.
- Les langues à satellites qui encodent la trajectoire dans un « satellite » associé au verbe, préfixe, préverbe, particule verbale (*satellite-framed language*) ex : les langues slaves, l'anglais, etc.
6. He **ran away** from the house.

2.3. *La notion de trajectoire*

Parmi les différentes composantes d'un événement spatial, la trajectoire est l'élément le plus fondamental. Son importance s'explique par le fait qu'il détermine le cadre spatio-temporel d'un tel événement. En effet, tout déplacement doit être vu en relation avec la temporalité.

Un déplacement implique que la figure se déplace sur l'axe de la trajectoire à partir d'une source (site initial) à un moment T1 pour atteindre un but (site final) à un moment T2. Pour aller d'un site à un autre, période relative à l'intervalle [T1- T2], la figure doit parcourir un trajet (site médian) qui est la portion d'espace située entre le point initial et le point final de la trajectoire. Le déplacement est ainsi conçu comme un déroulement dans un cadre temporel impliquant un début, un milieu et une fin dans lequel les différentes portions de l'espace initiale (source), médiane (trajet) et finale (but), réfèrent respectivement aux trois phases temporelles de l'événement, initiale, médiane et finale.

3. **Complexité typologique de l'arabe tunisien dans l'expression du déplacement**

Selon la typologie de Talmy, les langues sémitiques, et donc l'arabe tunisien, se définiraient comme des langues à cadre verbal. La particularité de telles langues est d'encoder la trajectoire du déplacement dans le verbe principal et d'exprimer la manière selon laquelle s'effectue le déplacement dans un autre syntagme associé. Cette affirmation pose d'emblée un certain nombre de problèmes : toutes les langues sémitiques présentent-elles les mêmes stratégies ? Et à l'intérieur des langues sémitiques, à quel « arabe » fait-on référence (littéral ou dialectal, mais dans ce cas, de quel pays, de quelle région ?) Peut-on parler d'une appartenance simultanée à différents types ? Ces questions ont d'autant plus de pertinence que l'on verra qu'on ne peut apparemment pas classer l'arabe tunisien dans les langues à cadre verbal. On verra que les stratégies que l'arabe tunisien présente pour l'expression de la trajectoire entrent mal dans la typologie de Talmy. En effet, on ne peut classer le tunisien comme langue à satellites car les préfixes verbaux n'expriment pas la trajectoire mais l'aspectualité ou la manière. On ne peut pas non plus classer cette langue comme langue à cadre verbal car elle recourt à d'autres stratégies. En effet, cette langue exprime la trajectoire à l'aide d'une « série » de deux verbes, l'un encodant la trajectoire et l'autre la manière (7). L'arabe tunisien utilise aussi une séquence verbe + participe actif (8), qui n'entre pas dans la catégorie des langues à cadre verbal selon Talmy. Enfin, nous verrons que le tunisien exprime également la trajectoire à l'aide d'un verbe seul (sans satellite ni syntagme associé). Il s'agit d'un phénomène de « fusion » qui n'entre pas dans la typologie de Talmy. Nous y reviendrons.

3.1. *L'arabe tunisien : langue à cadre verbal ?*

Nous allons voir que la réponse à cette question est non, car les stratégies employées sont autres. En effet, on trouve d'autres structures qui, on va le voir, n'entrent dans aucune des deux catégories de Talmy.

3.1.1. *Structure V1+V2*

7.	it-t ^s fol	dxæl	je-3r-i	li	l-madirsæ
	DEF-enfant	PAST.entrer.3MSG	PST.M-courir-3SG	à	DEF-école
	L'enfant est entré à l'école en courant.				

On a donc la structure lexicalisée suivante :

V1 [déplacement + trajectoire] +V2 [déplacement + manière].

Dans l'exemple (7), la trajectoire est exprimée dans le verbe '**dxæl**' 'entrer' qui décrit le passage de l'intérieur vers l'extérieur. La manière du mouvement, quant à elle, est exprimée

dans un autre verbe postposé au premier, le verbe ‘**jezri**’ ‘courir’ qui dénote l’aspect inaccompli (ou présent) du verbe par rapport à l’accompli (ou passé) du verbe ‘entrer’. En fait, le second verbe qui dénote la manière et qui est à l’inaccompli a une valeur atemporelle, ce qui révèle l’existence d’une hiérarchie entre les deux verbes (d’autres éléments le prouvent). Cette forme lexicalisée désigne un procès qui est parvenu à son terme (aspect perfectif). Il est important de souligner que cette structure syntaxique n’est pas attestée en arabe littéral, l’encodage de la manière se fait plutôt en littéral par le recours à ce qu’on appelle *nom d’action* ou « MASDAR ». C’est une dérivation à bornage aspectuel (ou procès achevé) du verbe courir qui n’est pas un verbe plein mais une sorte de nominalisation ayant également une valeur atemporelle.

3.1.2. Structure participe actif+ V

8. it-t ^ʕ fol	dæxil	je-3r-i	li	l-madisæ
DEF-enfant	PART.A.entrer.3MSG	PST.M-courir-3SG	à	DEF-école
L’enfant est en train d’entrer à l’école en courant.				

Ici, la trajectoire est encodée dans le participe actif (noté ici PART.A), tandis que la manière est encodée dans le verbe ‘courir’. Le participe actif, forme dérivée du verbe « entrer » dénote l’aspect **imperfectif** (non achevé) du procès. Le type de lexicalisation est le suivant :

PART.A [déplacement + trajectoire] + V [déplacement + manière].

La différence entre les deux types de structures lexicalisées (V1+V2 et participe actif+V) se situe au niveau aspectuel (perfectif vs imperfectif) et non pas au niveau spatial. En conséquence ces deux structures n’entrent pas dans la typologie de Talmy. Rappelons toutefois qu’il existe un lien très fort entre la notion d’espace et celle de temporalité. Cette typologie devrait pouvoir en rendre compte.

3.2. L’arabe tunisien : langue à satellites ?

Si on s’en tient à la typologie binaire de Talmy, si le tunisien n’est pas une langue à cadre verbal, ce devrait forcément être une langue à satellites. Si on trouve effectivement en tunisien des préfixes verbaux donnant des indications sur les différentes phases du déplacement (initial, médian, final), ils n’entrent pas dans le cadre de notre étude car ce sont avant tout des informations aspectuelles (procès en cours, procès achevé ou à venir) et ils ne donnent des informations sur le plan spatial qu’à un niveau dérivé. N’encodant pas la trajectoire ou le déplacement, ces préfixes n’entrent pas dans la catégorie des satellites de Talmy.

9a. it-t ^ʕ fol	je-mʃi	fi	e-ʃʃæriʃ
DEF-enfant	PST-M.marcher3SG	dans	DEF-rue
L’enfant marche dans la rue.			
9b. it-t ^ʕ fol	je-t-mæʃʃæ	fi	e-ʃʃæriʃ
DEF-enfant	PST-M.-manière-marcher3SG	dans	DEF-rue
L’enfant se promène lentement dans la rue.			

Ces préfixes encodent la manière (ex. *osciller* vs *voltiger*, *marcher* vs *se promener*, etc.) et non pas la trajectoire et n’entrent donc pas non plus dans la typologie de Talmy. Ce sont pourtant des structures qui encodent l’expression de l’espace dynamique et une typologie de ce domaine devrait pouvoir les intégrer. En conséquence, il semble qu’on ne puisse pas considérer le tunisien comme une langue à satellites.

3.3. « Fusion » des verbes de déplacement

Il existe en arabe tunisien des structures composées de deux verbes qui encodent la trajectoire + la manière. Ainsi, dans l’exemple suivant, le verbe tunisien ‘**fas^ʕ-it**’ associé à

la préposition ‘**mne**’ ‘de’ donne un sens que l’on pourrait traduire par ‘s’évader d’une manière discrète’. Cette structure encode le passage d’un espace à un autre (bouteille vers extérieur) ainsi que la manière selon laquelle le procès s’est déroulé (sortie de la grenouille de la bouteille). On peut parler dans ce cas d’une ‘fusion’ au sein du verbe [déplacement + trajectoire + manière] et d’une répartition de la trajectoire entre le verbe et la préposition ‘de’ qui indique la source.

10. i-33ranæ fas^oʕ-it mne e-ddæbbuzæ
 DEF-grenouille PAST.s’évader.discrètement-3FSG de DEF-bouteille
 La grenouille s’est évadée de la bouteille.

L’exemple suivant illustre également ce type de ‘fusion’ [déplacement + trajectoire + manière]. En effet, le verbe ‘**itsællit**’ exprime la manière selon laquelle la grenouille est sortie de la bouteille. La notion de trajectoire est répartie entre le verbe et la préposition.

11. i-33ranæ itsællil-t mne
 e-ddæbbuzæ
 DEF-grenouille PAS.esquiver.discrètement.avec difficulté-3FSG de
 DEF-bouteille
 La grenouille s’est esquivée et est sortie de la bouteille.

3.4. *Prédicats complexes ou « séries verbales »*

La série d’exemples que nous présentons ci-dessous met en lumière un autre type de lexicalisation qui n’est pas pris en compte dans la typologie binaire de Talmy, à savoir langue à cadre verbal, langue à satellites. En effet, il s’agit dans ces énoncés d’une juxtaposition de verbes qui ont le même sujet, l’un exprime la trajectoire ou le déplacement et l’autre la manière. Dès lors on se demande où on peut classer ces « prédicats complexes » et par conséquent l’arabe tunisien. Notons qu’à ce stade de notre étude, une analyse fine de ces structures reste à faire.

12. le-wlæjjed ʁræ3 i-duur fi i-33ardæ
 DEF-enfant PAST.sortir.3M SG PST.3M-tourner. SG dans DEF-jardin
 Le garçon est sorti se promener dans le jardin.

On a donc la structure suivante : **Prédicat complexe [trajectoire + déplacement]**.

Dans l’exemple ci-dessus, le verbe ‘sortir’ exprime le déplacement de l’intérieur vers l’extérieur, il spécifie le point initial et le point final de la trajectoire, avec changement de lieu. La trajectoire est donc dite « téléique », alors que le verbe ‘se promener’ ne précise pas le point final de la trajectoire, il réfère uniquement à une trajectoire médiane ou « atélique ». La figure est localisée dans un même lieu (l’enfant se promène dans le même endroit, il se déplace dans le jardin).

Cette distinction remet en question la notion de « motion event » de Talmy, du fait qu’on ne peut pas considérer le déplacement dans un endroit stationnaire (localisation) comme un déplacement qui lexicalise obligatoirement la trajectoire dans le verbe, dans le cas des *verb-framed languages*.

- 13a. it-t^ofol t^olaʕ je-zri fi i-ddru:3
 DEF-enfant PAST.monter.3MSG PST.3M-courir.SG dans DEF-escalier
 L’enfant a monté les escaliers en courant.
- 13b. it-t^ofol hbat^o i-næggiz mne i-ddru:3
 DEF-enfant PAST.descendre.3MSG PST.3M-sauter.SG de DEF-escalier
 L’enfant a descendu les escaliers en sautant.

On a donc : **SV [trajectoire + manière]**.

Notons que dans (13a) et (13b) la structure inverse SV [manière + trajectoire] est grammaticalement attestée dans un contexte particulier. Elle permettra de mettre l'accent sur la manière du déplacement et non sur la trajectoire.

- | | | | |
|---------------------------------------|---------------------|-------------------|-----|
| 14. it-t ^o fol | hbat ^s | je-mji | ʕlæ |
| sæq-i:h | | | |
| DEF-enfant | PAST.descendre.3MSG | PST.3M-marcher.SG | sur |
| pieds-POSS.3MSG | | | |
| L'enfant est descendu en marchant. | | | |
| SV [trajectoire + déplacement] | | | |
| 15. biʃ-ji-mji | | i-ʕu:m | |
| FUT-PST.3M-marcher.SG | | PST.3M-nager.SG | |
| Il va nager. | | | |
| SV [déplacement + mouvement] | | | |

Notons que dans (15), le verbe « marcher » postposé au préfixe « biʃ » marque du futur, est une construction figée qui correspond dans cet exemple à un procès envisagé par l'esprit.

L'interprétation purement spatiale passe en arrière-plan et le verbe « marcher » ne peut être considéré comme un verbe de déplacement. Le procès ne résulte pas de la sémantique du verbe « marcher » mais de ce qui émerge de sa combinatoire avec le verbe de mouvement « nager ». La trajectoire n'a donc aucun marquage spécifique, mais elle est inférée de l'interaction entre les deux verbes, « marcher » (à valeur aspectuelle) et « nager » (verbe de mouvement). Notons que cette construction figée ne s'emploie qu'avec des verbes de mouvement, tous deux encodant implicitement la trajectoire qui n'est pas linguistiquement exprimée mais doit être décodée à partir d'une représentation conceptuelle et un processus inférentiel. Comment peut-on donc rendre compte de la distribution de l'information spatiale dans le cas d'une interaction spatio-temporelle au sein de cette typologie qui ne prend en considération que les constructions syntaxiques permettant d'encoder la trajectoire ? Pris isolément, les éléments morphosyntaxiques ne rendent pas compte de la complexité d'une scène spatiale, d'où la nécessité de prendre en considération le sémantisme des différents éléments d'un événement spatial dans leur interaction et leur contexte particulier.

En ce qui concerne ces structures qui ressemblent à des « prédicats complexes » (et qu'il reste à étudier en détail), se pose le problème qu'elles n'entrent apparemment pas dans la typologie de l'expression du déplacement de Talmy. Slobin (2004) étend la typologie de Talmy en ajoutant les **langues T-équipollentes** (*equipollently-framed languages*), ensemble qui regroupe les langues à verbes sériels qui encodent la trajectoire et la manière « equipollently » dans deux verbes différents appartenant à une même clause, à radical double (verbes se composant d'un radical de manière et d'un radical de déplacement), à préverbes complexes (de manière et de trajectoire). Seulement et selon Fortis, (2007 : 3) « la pertinence du terme *equipollent* n'est pas toujours claire (il y a souvent un ordre fixe, avec des bases qui ont un statut différent). Les cas d'équipollence paraissent plus limités que Slobin ne l'entend (cas de double préverbe, manière et trajectoire ; cas de radicaux doubles etc.) ». D'après les exemples (13) et (14), l'arabe tunisien ne peut pas être classé selon ce troisième type de langues. En effet, le verbe qui exprime la manière est toujours placé après le verbe de déplacement dans le cas des « prédicats complexes » alors que dans les langues T-équipollentes le verbe de manière précède obligatoirement le verbe encodant la trajectoire. Cette dernière structure est attestée dans la langue, mais l'ordre des verbes n'est inversé que pour mettre l'accent sur la manière du déplacement dans un contexte particulier. D'où notre hésitation à les classer parmi les « séries verbales ».

4. Conclusion

Cette brève analyse typologique nous mène d'une part à remettre en cause le classement en bloc des langues sémitiques dans tel ou tel type. En ce qui concerne l'arabe tunisien, pour ce qui est de la lexicalisation des verbes de déplacement, il présente des exemples du type 1. On peut même dire qu'il est à dominance *verb-framed*, mais on ne peut le réduire à ce seul type de lexicalisation. Il existe d'autres stratégies que cette typologie ne prend pas en compte. En ce qui concerne le classement des langues à cadre verbal/satellites, il semble une fois de plus qu'on ne puisse classer toutes les langues sémitiques dans les langues à cadre verbal. En effet, il semble que ce ne soit pas le cas de l'arabe tunisien. On ne peut pas non plus, si on s'en tient strictement à la typologie de Talmy, classer l'arabe tunisien dans les langues à satellites, car les préfixes qui existent n'encodent pas la trajectoire.

Enfin, nous avons vu que certaines stratégies employées par l'arabe tunisien pour l'expression de la trajectoire n'entrent dans aucune des typologies, ni celle de Talmy, ni celle de Slobin, présentées brièvement dans cet article. Il s'agit de différents types de « prédicats complexes », d'une structure participe actif + verbe, de phénomènes de « fusion » et enfin de préfixes verbaux encodant la manière. Cela nous conduit en définitive à remettre en cause le classement en bloc des langues sémitiques et à nous interroger sur la validité de la typologie de Talmy en ce qui concerne la lexicalisation des verbes de déplacement et l'expression de la trajectoire (cadre verbal/satellites) ainsi que celle de Slobin (2004) qui est pertinente mais pas assez claire sur les langues « sérielles » ou à « prédicats complexes ». Bien qu'elle soit plus valide que celle de Talmy, la classification de Slobin laisse toutefois échapper certaines formes morphosyntaxiques présentes dans d'autres langues. L'arabe tunisien présente d'autres patterns qui ne rentrent pas dans le classement typologique des langues proposé par Talmy et Slobin. D'autres options permettent à la langue de distribuer l'information spatiale dans des participes, des prédicats complexes, des groupes prépositionnels juxtaposés ou coordonnés, etc. Les critères morphosyntaxiques ne suffisent pas à eux seuls pour rendre compte des différentes stratégies employées par la langue dans l'expression du déplacement. L'interaction entre la sémantique des différentes composantes d'un événement spatial et la pragmatique (inférence) semble être nécessaire pour une étude des stratégies typologiques du tunisien et des langues en général.

Références

- FORTIS, Jean-Michel. (2007). Typologie de l'expression linguistique du déplacement : un examen critique. *Séminaire Formes Symboliques*, 13 février 2007.
- SLOBIN, Dan.I. (2004). The many ways to search for a frog: linguistic typology and the Expression of Motion Events, in *Relating Events in Narrative: Typological and Contextual Perspectives*, vol. 2, STRÖMQVIST, Seven & VERHOEVEN, Ludo (Eds). Mahwah/New Jersey : Lawrence Erlbaum Associates, 219-257.
- TALMY, Leonard (1985). Lexicalization patterns: semantic structure in lexical form, in *Language Typology and syntactic description*, vol.3, SHOPEN, Timothy (Ed). Cambridge : Cambridge University Press, 57- 149.
- TALMY, Leonard (1991). Path to realization: a typology of event conflation. *Proceedings of the Berkeley Linguistics Society* 17, 480-520. [modifié dans 2000, tome 2 : 213-287]
- TALMY, Leonard (2000). *Toward a cognitive semantics*, vol.2, Cambridge, Mass. : M.I.T. Press.

LA MOTIVATION DANS LES NOMS DES MALADIES EN ROUMAIN : DES EXEMPLES DE TABOU LINGUISTIQUE

Carmen SCARLAT

carmen.scarlat@u-grenoble3.fr

Centre de Dialectologie Université Stendhal Grenoble 3
38040 Grenoble cedex 9

Abstract : This study is based on a corpus with different sources: Romanian linguistic atlases, unpublished material coming from dialectal surveys from the beginning of the 20th century, personal surveys (2002-2004), dialectal glossaries and monographs. The aim is to illustrate the principles underlying the naming process in the field of names for diseases. Special attention is paid to the existence and the predominance of linguistic taboos.

Notre étude est une analyse lexicale et sémantique des désignations dialectales des maladies dans le domaine daco-roumain, dans le but de mettre en évidence les mécanismes présidant à la formation de ces noms, les procédés cognitifs et linguistiques sous-jacents à leur choix. Nous nous inscrivons en premier lieu dans la démarche initiée par les deux atlas de 2^e génération, *ALE et aliR*, qui apportent comme nouveauté par rapport aux autres atlas l'interprétation des données brutes. Les questions que nous nous posons sont les suivantes : Comment se fait le processus de dénomination ? Quels rapports, s'ils existent, gèrent la relation entre le terme et la réalité qu'il est censé représenter ?

Le principe de la motivation¹ est exploré par la linguistique pour ainsi dire dès ses débuts. De Saussure (1967) examine déjà le domaine du motivé par rapport à la définition du signe linguistique. De manière implicite, le principe de motivation est, comme l'a souligné Ovtchinnikova (2004 : 5), « un des fondements de toute recherche étymologique, dans la mesure où l'étymologie prend en compte l'exigence de continuité relationnelle dans l'histoire du mot ». Guiraud (1986) a fait de la motivation une des pierres angulaires de l'analyse des structures sémantiques de la langue. De même, depuis quelque temps, on voit avancer des approches de recherche linguistique qui font de la motivation lexicale une matière d'étude à part entière. On assiste ainsi à la publication d'atlas interprétatifs, d'études de vocabulaire² ou de thèses³ spécialement orientées vers la représentation du lexique dans sa dimension motivationnelle.

Une conception convergente à celle de Guiraud (1986) a été formulée par Alinei (1979) : celui-ci envisage une structure dynamique du signe qui prend en compte sa motivation et une structure statique du signe dans laquelle la motivation n'intervient pas. Alinei a le mérite

¹ La motivation est généralement définie en linguistique en termes de « relation de nécessité » (Dubois *et al.*, 1973 : 328) qui s'établit entre la forme du signifiant et le signifié. Par extension, le principe de motivation peut être pris dans l'optique du pronostic sur sa signifiante à partir de la forme signifiante. Toute forme linguistique dont la composition morphologique est facilement identifiable ou dont le procédé de création sémantique est transparent peut être jugée motivée. Réciproquement, la motivation peut être comprise en termes de nécessité de choix de la forme signifiante en fonction du signifié à exprimer. Autrement dit, la sélection du matériau linguistique est motivée par les besoins de l'expression dénomminative.

² Cf. notamment les articles parus dans les publications de revues telles que *Géolinguistique* à Grenoble, *Quaderni di Semantica* à Bologne, etc.

³ Cf. Kutangidiku (1995), Ovtchinnikova (2004), Signorini (2005).

d'être à l'origine d'une réflexion théorique, formulée explicitement, sur la motivation à l'origine de la création lexicale et sur la cartographie motivationnelle. Cela n'avait pas été fait auparavant, même si la motivation lexico-sémantique a été rapidement individualisée comme une composante importante, sinon fondamentale, de l'analyse étymologique, et mise à la base de la distinction non négligeable et qui existe depuis longtemps entre lexèmes transparents et lexèmes opaques. Alinei (1997 : 3) va plus loin, en proposant de faire de la motivation une sous-discipline indépendante de la linguistique et de lui donner l'étiquette terminologique, moins ambiguë, d'*iconomastique*⁴. Le chercheur revient, dans ses travaux, à maintes reprises, sur le rôle et le positionnement du motif dans le fonctionnement linguistique et son analyse est basée justement sur la dualité entre l'arbitraire et le motivé.

Notre hypothèse de travail est que le système des noms des maladies constitue, tout comme celui des plantes et des animaux, des termes de couleur et des systèmes de parenté, un domaine structuré, voire même hiérarchisé. Nous avons procédé à un « décortilage » de l'ensemble des dénominations relevées afin d'en identifier les « raisons », les bases cognitives se trouvant derrière la création lexicale qui n'est nullement un fruit du hasard. Ensuite, nous avons réduit la grande variété des appellations à plusieurs types motivationnels ou onomasiologiques. Pour le présent article, nous allons nous intéresser plus particulièrement, à l'intérieur de cette classification, à l'existence et à la prépondérance du phénomène du tabou linguistique, à sa survivance jusqu'à nos jours.

Notre corpus se caractérise par une grande hétérogénéité des sources : enquêtes personnelles (nous avons réalisé, entre 2002 et 2004, 34 enquêtes sur le territoire de la Roumanie), atlas linguistiques roumains, données inédites restées manuscrites qui proviennent des enquêtes du début du XX^e siècle, dictionnaires, glossaires et monographies. Le corpus relève ainsi d'époques différentes, allant depuis la fin du XIX^e siècle jusqu'à nos jours, ce qui nous permet d'opérer des analyses diachroniques et diatopiques.

1. Présentation et analyse des résultats

L'analyse de la façon dont le daco-roumain a adopté et développé le système populaire de dénomination des maladies que nous avons relevées nous a permis d'identifier plusieurs classes motivationnelles, plusieurs facteurs qui sont pris en compte lorsqu'il s'agit de nommer une maladie : 1) l'agent pathogène (réel ou imaginaire) de la maladie ; certaines espèces d'animaux (papillons, chenilles, vers, crapauds, etc.) sont des éléments qui, par leur installation dans l'organisme humain, provoquent, dans l'imaginaire populaire (mais pas seulement), l'état de maladie (Purdela Sitaru, 1999 : 170-171) : inflammations et maladies de l'épiderme (ressemblance enflure – animal, comme *gușter* 'lézard vert' et 'croup', *curu găinii* 'le cul de la poule' et 'dartres'), psychoses légères (vers, insectes, oiseaux), maladies graves ('chien', 'écrevisse', pour la rage et la cancer respectivement) ; 2) les parties du corps humain atteintes par la maladie ; le nom s'explique ici à partir du trait de la localisation de la maladie : *vintre* 'ventre, estomac' et 'diarrhée'... ; 3) la catégorie des personnes susceptibles d'avoir/transmettre la maladie : *boala turcilor* 'la maladie des Turcs' et 'peste' ; *boală franțuzească* 'maladie française' et 'syphilis' ; *boală englezească* 'maladie anglaise' et 'rachitisme'... ; 4) les symptômes sous lesquels se manifeste la maladie ; c'est le groupe de motivations le plus riche et le plus varié : *pârțică* 'petit pet', 'diarrhée' ; *căimăceală* 'épilepsie' < *caimac* 'crème (du lait bouilli)'... ; couleurs : *flori roșii* 'fleurs rouges', *poală roșie* 'jupe rouge' pour 'menstruation'... ; 5) le caractère de la maladie : gravité, durée, taille, généralement des enflures, comme *boala lungă* 'la maladie longue' et 'fièvre typhoïde' ; *bube mari* 'boutons grands', pour la 'variole'... ; 6) le désir de chasser la

⁴ Il y a équivalence entre iconicité et motivation, et *icône* est pris dans le sens terminologique de Peirce (1978).

maladie : *cel pierit* ‘celui disparu’, *perit negru* ‘disparu noir’ pour la ‘syphilis’... ; 7) le moyen de traitement : *argint viu* ‘argent vif’ et ‘syphilis’.

À côté de ces désignations à motivation transparente, il y a également des emprunts, des désignations à motivation perdue, à motivation inconnue. Nous allons analyser uniquement les désignations qui peuvent être interprétées comme des substituts euphémiques.

2. La place des désignations tabouistiques

Ce qui frappe tout d’abord est la proportion très importante des désignations motivées, qui atteigne les 85%, dont les 53% sont occupés par les désignations tabouistiques et les 15% restants représentent des appellations dont la motivation n’est pas perçue par les locuteurs du roumain.

3. Tabou linguistique : causes possibles et types de maladies touchées

De nombreuses études (Meillet, 1921 ; Zelenin, 1988, 1989a et b ; Bonfante, 1939 ; Bally, 1951 ; Galli de’ Paratesi, 1964 ; Benveniste, 1966, 1974 ; Alinei, 1997, 2004, ou, dans le domaine roumain, Iordan, 1962 ; Candrea, 1999 ; Purdela Sitaru, 1999) ont traité des multiples raisons de cette interdiction. Beaucoup d’entre elles sont valables pour tous les substituts euphémiques, comme le désir d’éviter la vérité désagréable que représente l’état de maladie, la peur superstitieuse d’attraper la maladie, selon la croyance que le fait de prononcer le vrai nom d’une maladie signifie la faire venir, croyance encore vivante sur le territoire de la Roumanie, comme nous avons pu le constater dans nos enquêtes et analyses de corpus.

Ensuite, d’autres facteurs peuvent intervenir, spécifiques aux différents groupes de maladies qui entrent en jeu. Parmi les maladies dont le nom est le plus frappé d’interdiction on compte, selon les chercheurs qui se sont penchés sur la question, les maladies sexuelles, les maladies de la peau, les maladies perçues comme dangereuses et celles mortelles. Cela se vérifie également dans notre corpus. Dans le cas des maladies vénériennes, le tabou est lié à la répression sexuelle et à la pudeur. Avant d’être graves, ces maladies sont honteuses et leur gravité est augmentée par les superstitions et le sens de la culpabilité. L’interdiction du nom des maladies de l’épiderme est probablement à mettre en rapport avec le dégoût devant ce qui est répugnant à voir. Les maladies redoutables forment un autre groupe dont les appellations sont tabouisées. Les recherches antérieures ont montré que le mal qu’on craint le plus semble être le cancer ; l’épilepsie est, elle aussi, chez de nombreux peuples, un élément de contenu préférentiel de l’interdiction linguistique (Purdela Sitaru, 1999 : 165), à côté des maladies mentales graves, probablement en raison de leurs formes de manifestation, qui font penser à l’intervention des esprits malins. D’autres catégories de maladies dangereuses sont celles de l’appareil digestif, comme la diarrhée, spécialement chez les enfants en bas âge (chez lesquels elle était une cause fréquente de mortalité) et les névralgies, telles que la paralysie, totale ou partielle.

4. Typologie du tabou linguistique

Nous pouvons regrouper en trois grandes catégories les moyens de création des *remplaçants* de toutes ces maladies dont la prononciation est frappée d’interdiction : les désignations propitiatoires, les désignations allusives et les termes offensifs et ironiques.

4.1. Les désignations propitiatoires

Elles sont le résultat de l’essai de *captatio benevolentiae* des forces infernales, dont les démons des maladies, qui peuplent l’univers et qui habitent le corps de la personne malade, dans une conception animiste du monde. Ce sont des appellations cryptiques, qui expriment

l'affection et le respect, dans le but de flatter la maladie et d'inverser ses effets négatifs, d'arriver à la faire quitter le corps dont elle a pris possession. Elles sont pour la plupart des appellations anthropomorphiques (adjectifs à connotation positive, généralement substantivisés), qui témoignent largement de l'image personnifiée qu'ont les maladies dans l'esprit humain.

Parmi les termes affectueux, un groupe lexical distinct est représenté par toute une série d'adjectifs, en emploi substantivisé, avec un déterminant défini ou un démonstratif, qui constituent les noms donnés en daco-roumain aux fées méchantes et à plusieurs maladies redoutables telles que la paralysie, la congestion cérébrale, la névralgie, l'épilepsie, la 'blennorragie' : [tʃ'ele kur'ate] 'les propres', [frumw'asele] 'les belles'⁵, [milost'ivele] 'les charitables' ; [bl'inda] 'l'aimable' pour l'urticaire, etc. Ces fées ressemblent, la plupart du temps, à des jeunes filles, belles, vêtues de blanc, qui apparaissent la nuit et qu'on « redoute parce qu'elles aiment les jeunes gens qu'elles veulent posséder, mais leur passion rend ceux-ci malades » (Taloş, 2002 : 87). Leurs appellations relèvent toutes du phénomène du tabou linguistique, à l'origine duquel se trouvent la peur inspirée par ces êtres dangereux, combinée avec la croyance dans le pouvoir magique du mot, le tout se prolongeant dans la crainte des noms qui leur sont liés.

D'autres adjectifs déterminent le nom d'un état pathologique (ce que nous avons appelé 'génériques' dans la taxinomie des maladies, pour les distinguer des hypergénériques, de sens évusif) comme *dulce* 'doux, sucré', dans [b'ube dultʃi] 'boutons doux', pour désigner l'eczéma, *frumos* 'beau', comme dans [b'ubə frumw'asə] 'bouton beau' et 'anthrax'.

On compte aussi des adjectifs de couleur, tels que [alb] 'blanc', dans [b'ube 'albe] 'boutons blancs', pour 'dermatites' et [vərs'at alb] 'variole blanche' pour 'rougeole'. Le lien avec la couleur n'est pas seulement absent, mais il représente l'opposé de la situation réelle, désagréable, qui est ici rendue par un adjectif connoté positivement.

C'est probablement une conséquence du même phénomène qui s'est matérialisée dans l'apparition de faux diminutifs formés à partir des noms de maladies : [trinzi'ori] de [trinzi] 'hémorroïdes', avec le suffixe diminutif composé *-işor*, etc. Les noms des démons sont souvent des diminutifs, comme [mətuʃ'ika] 'la petite tante', qui désigne aussi la 'malaria'.

Dans la catégorie des termes respectueux, on peut mentionner les adjectifs épithètes et les noms de parenté (mère, tante), issus probablement de la crainte inspirée par la maladie, qui apparaît ainsi respectivement sacrée ou anthropomorphisée : [tʃele sf'inte] 'les sacrées', pour 'fées méchantes' et 'blennorragie'; [bw'ala sf'intə] 'la maladie sacrée' pour 'dysenterie', d'un côté, et [m'aika bətr'inə] 'la mère vieille', [m'aika k'aləa] 'la mère chemin', [m'aika kələtw'arəa] 'la mère voyageuse', pour la 'peste ou le choléra des gallinacés', de l'autre côté.

4.2. Les désignations allusives

Elles expriment le contournement de la réalité désagréable que représente l'état de la maladie ; pour ce faire, plusieurs moyens sont mis en œuvre, que nous avons regroupés en : a) substituts pronominaux et déterminants, qui peuvent être des pronoms personnels ou démonstratifs : le pronom sujet [j'ele] et une forme de politesse [d'insele] 'elles' pour les 'fées méchantes' et de nombreuses maladies, 'rhumatisme', 'spondylose', 'épilepsie' ; ['alta]

⁵ En allemand, par exemple, dans la même attitude de *captatio benevolentiae*, on appelle l'érysipèle *das Schöne* 'la belle' (Augusto & Melka, 1996 : 63).

‘l’autre’, [‘alte ‘alɛa] ‘d’autres (choses)’, ‘celles-là’ pour l’‘épilepsie’ ; ces exemples représentent des cas de généralisation maximale. Leur emploi pourrait s’expliquer par le désir de ne pas attirer l’attention d’un esprit ou de mettre en colère celui qui occupe le corps humain ; b) hyperonymes/(hyper)génériques de la taxinomie des maladies : [bw‘alə] ‘maladie’ pour la ‘fièvre typhoïde’ ; [a bw‘alə] ‘cette maladie-là’, [nek‘az] ‘difficulté’ pour l’‘épilepsie’ ; avec ces appellations, le degré de généralisation se rétrécit. On entre dans le champ des maladies, par le haut de la hiérarchie, la catégorie superordonnée, la plus englobante. Les génériques fonctionnent, dans le lexique de la médecine populaire, sur le même schéma que les hypergénériques : [ab‘ubə] ‘ce furoncle-là’, pour ‘anthrax’, ‘abcès alvéolaire’ ; c) (hyper)génériques suivis de spécificateurs plus ou moins évasifs, qui se trouvent un rang plus bas : [r‘əwu bəj‘etsilor] ‘le mal des garçons’ et [bw‘ala kopij‘ilor] ‘la maladie des enfants’ pour l’‘épilepsie’ ; [b‘ube mitʃi] ‘boutons petits’, pour la ‘rougeole’ et la ‘scarlatine’ ; d) autres substituts : des plus généraux [l‘ukru] ‘chose’, [tʃɛas] ‘moment’, à ceux qui le sont moins ; dans ce dernier cas il s’agit des supposées causes externes des maladies dont le nom est frappé de tabou, notamment le soleil, le vent, la pluie, qui figurent dans les désignations de plusieurs maladies. D’après la pensée populaire, les phénomènes naturels sont la plupart du temps des esprits néfastes (Augusto & Melka, 1996 : 58)⁶ : [‘ale din vint] ‘celles du vent’, et ‘paralysie’, ‘rhumatisme’, ‘spondylose’, ‘blennorrhagie’ ; [vint rəu] ‘vent mauvais’ et [vint sek] ‘vent sec’ pour ‘congestion cérébrale’, ‘paralysie faciale’, ‘névralgie’ ; [bl‘inde din plw‘aje] ‘douces de la pluie’ et ‘urticaire’, etc.

4.3. Les termes offensifs et ironiques

Ce sont des noms ‘moches’, pour insulter la maladie, utilisés lorsque les autres moyens n’ont pas abouti. La peur de prononcer le nom d’une maladie, dont nous avons parlé ci-dessus, s’actualise ainsi de façon tout à fait contraire : une sorte de courage conduit à chasser le mal en utilisant ces noms insultants, parmi lesquels nous avons identifié plusieurs catégories : a) adjectifs offensifs : [spurk] ‘sale’ et ses dérivés [spurk‘at], [spurkətʃ‘une] pour ‘furoncle’, ‘anthrax’, ‘impétigo’ ‘diarrhée’ (principalement des maladies de l’épiderme) ; [rəu] ‘mauvais, méchant’ et son dérivé [tʃɛas rəu] ‘moment mauvais’, pour ‘épilepsie’ et ‘tachycardie’ ; [l‘ukru rəu] ‘chose méchante’ pour ‘épilepsie’ ; [vrəzm‘afa] ‘l’ennemie’ pour ‘anthrax’⁷ ; b) adjectifs de couleur à connotation sombre, négative : ‘noir’ qui entre dans la composition de plusieurs noms de maladies là où il ne traduit pas l’aspect physique de l’affection, comme dans [bw‘ala ɲɛ‘agrə] ‘la maladie noire’ pour ‘épilepsie’ ; c) imprécations : ce sont des syntagmes, des formulettes, issus des textes des incantations ; la plupart sont les constructions avec des verbes au conjonctif daco-roumain, qui font penser aux imprécations ; elles font également référence à des maladies fort redoutables dans la conscience populaire : [d‘ukəsenv‘int] ‘qu’il s’en aille dans le vent’, [m‘intʃel k‘oka] ‘que le

⁶ Dans les langues romanes, nous avons trouvé des noms similaires en portugais, où le glaucome est appelé *lunar*, dérivé de *luna*, l’insolation *mal do sol* ‘mal du soleil’, l’eczéma *orvalhada*, et une maladie vénérienne *orvalhico*, de *orvalho* ‘rosée’ (Augusto & Melka, 1996 : 58). En albanais et dans les langues slaves les exemples sont nombreux : en albanais, on nomme l’épilepsie *ajó e tokës* ‘celle de la terre’ ; en bulgare, *vânkašna bolest* ‘la maladie du dehors’ ; en serbe, la même maladie est appelée *gorska bolest* ‘la maladie de montagne’ (Candrea, 1999 : 272).

⁷ Ces types lexicaux se retrouvent dans de nombreuses langues d’Europe (en portugais, l’*érysipèle* est appelé *espírito mau* ‘esprit mauvais’, l’*anthrax* *maldita* ‘maudite’, en italien l’épilepsie s’appelle, entre autres, *bruto male* ‘mal mauvais’, etc. (Augusto & Melka, 1996 : 64 ; Candrea, 1999 : 272).

grumeau le mange' pour 'diarrhée' ; s'y ajoutent aussi les désignations formées à partir du verbe 'disparaître', [tʃel pjer'it] 'celui disparu', pour 'syphilis' ; d) termes ironiques, assez peu représentés dans notre corpus, au nombre de 3 : [domn'ije] 'princerie' et [fudul'ije] 'fierté' pour 'gale', [bojer'ije] 'noblesse' pour la 'syphilis'.

5. Quelques considérations étymologiques

Sur l'ensemble des noms de maladies que nous avons recueillis, une analyse étymologique met en évidence l'existence de sources diverses et plus ou moins attendues. Une partie des éléments lexicaux sont hérités du latin, les autres sont des emprunts au slave⁸, ture, hongrois et allemand. Il est intéressant de noter l'absence, en daco-roumain standard, des mots d'origine latine pour 'maladie' et 'malade', qui sont des emprunts au slave (*boală, bolnav*). Cela pourrait s'expliquer par le même phénomène de tabou, alors que les termes exprimant les notions contraires, *sănătate* 'santé', *sănătos* 'en bonne santé', n'étant pas frappés d'interdiction, se sont gardés.

6. La géosynonymie des termes du vocabulaire de la médecine populaire

La géosynonymie est un des traits facilement observables à propos de notre corpus. Par exemple, pour désigner l'épilepsie, les parlers daco-roumains connaissent une grande variation lexicale. C'est la maladie la plus redoutée, en raison de l'intervention des mauvais esprits, d'où la grande majorité de termes tabouistiques qui la désignent : *baiul cel rău* 'le mal méchant', *betegeala rea*, *boala rea* 'la maladie méchante', *ceas rău* 'moment mauvais', *lucru rău* 'chose mauvaise', *nevoia rea* 'difficulté mauvaise', *ânevoie* 'cette difficulté-là', *nevoie* 'difficulté', *baiu cela* 'cet ennui-là', *boala aia*, *aboală* 'cette maladie-là', *răul acela* 'ce mal-là' etc. On compte 38 appellations, qui font de l'épilepsie la maladie avec le degré de généralisation maximale.

Le revers de la situation précédente est facilement observable : le même syntagme finit par être employé pour désigner plusieurs maladies, ce qui peut soulever des difficultés au niveau de l'identification correcte du sens, lorsque la construction en question est décontextualisée. Parmi les formations les plus chargées sémantiquement, on peut rappeler : *boala rea* 'la maladie mauvaise' : 1. 'maladie incurable' ; 2. 'épilepsie' ; 3. 'fièvre typhoïde' ; 4. 'apoplexie' ; 5. 'rage' ; *buba rea* 'le bouton mauvais' : 1. 'anthrax' ; 2. 'érysipèle' ; 3. 'cancer' ; 4. 'adénopathie' ; 5. 'panaris', etc.

7. Ouverture vers le cognitif

Si on compare les résultats de notre étude à ceux qui ont été obtenus dans d'autres domaines linguistiques (Bonfante, 1939 pour l'italien et autres langues indoeuropéennes ; Havers, 1946 pour l'allemand ; Ruud, 1960 pour le malgache ; Zelenin, 1988, 1989a et b, pour le slave ; Candrea, 1999, qui, en plus du vocabulaire roumain, donne de nombreux exemples tirés des langues slaves, du hongrois, albanais, ture, grec, etc. ; Augusto & Melka, 1996, pour le portugais et le français, etc.), on peut se demander si les similitudes qui se profilent ne témoignent pas d'une résurgence « spontanée » de tel ou tel schéma de création. La vision conceptuelle de l'être humain sur les maladies semble suivre, comme l'a souligné la linguistique cognitive, des principes fondamentaux qui se manifestent indépendamment

⁸ Les traits communs de l'imaginaire populaire daco-roumain et celui slave (Zelenin, 1988, 1989a et b) se reflètent largement au niveau du vocabulaire, non seulement au niveau quantitatif des emprunts, mais surtout du point de vue qualitatif, des évolutions et des structures sémantiques similaires des termes daco-roumains d'origine non slave et les mots slaves correspondants (Purdela Sitaru, 1999 : 92).

dans les langues (Berlin, 1992). Le champ sémantique étudié apparaît ainsi structuré sur la base des mêmes principes qui sous-tendent le développement cognitif et culturel (Atran, 1998), et qui sont nécessairement logiques, physiques, utilitaires, sociaux, etc. La maladie en tant que référent-objet d'une dénomination touche l'homme de très près. Le malaise causé par l'absence de santé, la méconnaissance de l'origine du mal et du remède ou encore l'indéfinition nosologique surtout des maux intérieurs, dermiques ou nerveux, rendent l'individu peureux et conscient de son impuissance envers la maladie.

Ces sentiments conduisent alors souvent à attribuer le mal à des forces supérieures, à des animaux, à des phénomènes naturels et à avoir recours à des noms *noa*, utilisés pour éviter le nom réel, à cause du tabou (Alinei, 2004 : 113). De plus, les mêmes moyens linguistiques qui témoignent de cette interdiction se retrouvent dans d'autres champs lexicaux, comme dans celui des animaux ou, dans une moindre mesure, des plantes (Alinei, 1997 ; Scarlat, 2006 et en préparation).

Comparé à d'autres sociétés de l'Europe occidentale qui ont perdu petit à petit leur savoir populaire, le roumain semble avoir gardé, dans son vocabulaire de la médecine populaire mais pas seulement, des traces de ce savoir. De nombreux termes de notre corpus sont effectivement entrés dans la langue standard ; une partie non négligeable (17%) des désignations de notre corpus a été recueillie lors de nos enquêtes personnelles. À l'intérieur de ces dénominations issues de nos enquêtes, la plupart des noms de maladies (83%) cachent une motivation. Dans des cas très rares les désignations sont le résultat d'accidents linguistiques. Celles qui résultent d'une déformation de la dénomination primaire (comme pour les attractions paronymiques) ne sont néanmoins pas dues au hasard : derrière ces mutations il y a une motivation. Cela confirme notre hypothèse de départ : il y a bien, en roumain, des conceptualisations particulières touchant aux maladies, i.e. des modes de représentation populaire qui, de plus, viennent confirmer les résultats obtenus dans d'autres domaines linguistiques. Parmi ces bases cognitives, le phénomène du tabou linguistique est certes moins productif par rapport aux périodes antérieures (dans le cadre des désignations provenant des autres sources, les noms tabouistiques sont de loin majoritaires), mais il continue à être à la base d'une grande partie de ce lexique populaire, dialectal (40%), dont de nombreux termes sont entrés en roumain littéraire.

Références bibliographiques

ALE = (1983-). *Atlas Linguarum Europae*. Assen : Van Gorcum, Roma : Istituto Poligrafico e Zecca dello Stato.

ALINEI, Mario (1979). The structure of meaning, in *A semiotic Landscape*, Chatman, Seymour, Eco, Umberto, Klinkenberg, Jean-Marie (Eds). Berlin : De Gruyter-Mouton, 499-503.

- (1997). Magico-religious motivations in European dialects : a contribution to archaeolinguistics, *Dialectologica et Geolinguistica*, 5 : 3-30.

- (2004). The Role of Motivation ("iconymy") in Naming : Six responses to a List of Questions, in *Nature knowledge. Ethnoscience, Cognition and Utility*, Sanga, Glauco & Ortalli, Gherardo (Eds). Venezia : Istituto Veneto di Scienze, lettere ed Arti, 108-118.

AliR = (1996-). *Atlas Linguistique Roman*. Roma : Istituto Poligrafico e Zecca dello Stato.

ATRAN, Scott (1998). Folkbiology and the anthropology of science : Cognitive universals and cultural particulars, *Behavioral and Brain Sciences*, 21 : 547-609.

AUGUSTO, Maria Celeste & MELKA, Francine (1996). Motivations linguistique et ethno-linguistique : le cas des dénominations populaires des maladies en français et en portugais, *Bollettino dell'Atlante Linguistico Italiano*, III, 20 : 49-77.

BALLY, Charles (1951). *Traité de stylistique française*. Paris : Klincksieck.

BENVENISTE, Émile (1966 ; 1974). *Problèmes de linguistique générale* (2 vol.). Paris : Gallimard.

- BERLIN, Brent (1992). *Ethnobiological classification : principles of categorization of plants and animals in traditional societies*. Princeton : Princeton University Press.
- BONFANTE, Giuliano (1939). Études sur le tabou dans les langues indo-européennes, in *Mélanges de linguistique offerts à Charles Bally*. Genève : Université de Genève, Faculté des Lettres.
- CANDREA, Ion-Aurel (1999). *Folclorul medical român comparat*. București : Polirom.
- DE SAUSSURE, Ferdinand (1967). *Cours de linguistique générale*. Paris : Payot.
- DUBOIS, Jean, GIACOMO, Mathée, GUESPIN, Louis, MARCELLESI, Christiane, MARCELLESI, Jean-Baptiste & MÉVEL, Jean-Pierre (1973). *Dictionnaire de linguistique*. Paris : Larousse.
- GALLI de' PARATESI, Nora (1964). *Semantica dell'eufemismo*. Torino : G. Giappichelli.
- GRAUR, Alexandru (1963). *Etimologii românești*. București : Editura Academiei.
- HAVERS, Wilhelm (1946). *Neuere Literatur zum Sprachtabu*. Wien : Akademie der Wissenschaften.
- GUIRAUD, Pierre (1986). *Structures étymologiques du lexique français*. Paris : Payot.
- JORDAN, Iorgu (1962). *Lingvistica romanică. Evoluție. Curente. Metode*. București : Editura Academiei Republicii Populare Române.
- KUTANGIDIKU, Tshimanga (1995). *La motivation dans la création lexicale des noms de petits animaux chez les Bantu du Zaïre*. Thèse de doctorat préparée sous la direction de Michel Contini. Grenoble : Université Stendhal-Grenoble 3.
- MEILLET, Antoine (1921). Quelques hypothèses sur les interdictions de vocabulaire dans les langues indo-européennes, in *Linguistique historique et linguistique générale I*. Paris : Champion, 281-291.
- OVTCHINNIKOVA, Ekaterina (2004). *Les mots de la cuisine. Lexique de la cuisine familiale dans les parlers d'oc. Problèmes de sémantique, d'étymologie et de géographie linguistique*. Thèse de doctorat préparée sous la direction de Jean-Philippe Dalbéra. Nice : Université de Nice-Sophia Antipolis.
- PEIRCE, Charles (1978). *Écrits sur le signe*. Paris : Éditions du Seuil.
- PETROVICI, Emil (1956-1972). *Atlasul Lingvistic Român* (7 vol.). București : Editura Academiei Române.
- POP, Sever (1938). *Atlasul Lingvistic Român I, I*. Cluj : Muzeul limbii române.
- PURDELA SITARU, Maria (1999). *Etnomedicină lingvistică*. Timișoara : Amarcord.
- RIEGLER, Richard (1933). Tabunamen des Fuchses, *Archivum Romanicum*, XVII : 405-409.
- RUUD, Jørgen (1960). *Taboo. A study of Malagasy Customs and Beliefs*. Oslo : Oslo University Press.
- SCARLAT, Carmen (2006). Les noms des maladies dans les noms des plantes: quelle palce pour le tabou linguistique. Le cas du daco-roumain, *Quaderni di semantica*, 1-2, *Studi in onore di Mario Alinei* : 391-405.
- (en préparation). *Étude étymologique, sémantique et cognitive des phytonymes dans le domaine daco-roumain*. Thèse de doctorat préparée sous la direction de Jeanine Médélice. Grenoble : Université Stendhal Grenoble 3.
- SIGNORINI, Céline (2005). *La motivation sémantique dans la création lexicale : les phytonymes de l'Arc Alpin*. Thèse de doctorat sous la direction de Michel Contini. Grenoble : Université Stendhal Grenoble 3.
- TALOȘ, Ion (2002). *Petit dictionnaire de mythologie populaire roumaine*. Grenoble : ELLUG.
- ZELENIN, Dimitrij (1988). Tabù linguistici nelle popolazioni dell'Europa orientale e dell'Asia settentrionale, *Quaderni di semantica*, IX, 2 : 187-317 ; (1989a), X, 1 : 123-180 ; (1989b), X, 2 : 183-276.

PHONETIC ALIGNMENT AND THE PHRASE TONE / ‘TRAILING’ TONE AMBIGUITY IN IRISH ENGLISH

Jennifer SULLIVAN

Jennifer.sullivan@ucd.ie

School of Irish, Celtic Studies, Irish Folklore and Linguistics

University College Dublin, Dublin 4

Irlande

Abstract : This paper explores the ambiguity between phrase tones and ‘trailing’ tones in Irish English nuclear falls. Data is examined in the context of the Constant Interval and Segmental Anchoring Hypotheses. Results show an ambiguity in that the L tone in two dialects did not appear to behave strictly as either a phrase tone or a ‘trailing’ tone, which has consequences for the annotation of these dialects.

1. Introduction

1.1. *Phrase tones and ‘trailing’ tones*

The notion of the phrase tone originated in Bruce’s seminal work on Swedish (1977) and was incorporated into Pierrehumbert’s similarly influential (1980) work on American English. Phrase tones in Pierrehumbert (1980) were believed to lie somewhere between the nuclear (last) pitch accent and the Intonational Phrase (IP) boundary, possibly at the edge of the accented word. However, their exact location and alignment was not very clearly defined. With the advent of Beckman and Pierrehumbert (1986), the phrase tone acquired a new function of being the delimiter of the Intermediate phrase and this function was retained in the development of ToBI (Tone and Break Indices) (Silverman *et al.*, 1992). More recently, Grice, Ladd and Arvaniti (2000) and Lickley, Schepman and Ladd (2005) have presented evidence in favour of the first postnuclear stressed syllable being the location of the phrase tone, a change from Pierrehumbert’s original suggestion.

Since the arrival of ToBI, numerous other systems have been developed, including the IViE system (Intonational Variation in English) (Grabe, Nolan & Post, 2003). Although basically similar to ToBI, IViE has a number of key differences, most important among them (for this paper) being that IViE omits phrase tones and is by no means the only system to do so. Indeed IViE is based more so on the works like the one of Gussenhoven (1999) (who also omits phrase tones) than on the ToBI system itself. The difference between ToBI and IViE in terms of their inclusion/omission of phrase tones has relevance for this paper, as these two systems would annotate nuclear falls somewhat differently from each other. The ToBI system would annotate such a fall as H* on the accented syllable followed by L- (phrase tone), followed by L%. The IViE system, on the other hand, would annotate a nuclear fall rather as H*L (‘trailing’ tone) with an optional boundary tone. A ‘trailing’ tone has long been associated with the notion of a fixed distance between it and the ‘starred’ tone of the pitch accent (Pierrehumbert, 1980). This notion of a fixed distance has persisted up to this day in ToBI-based analyses, despite a growing body of evidence which suggests that it may not mirror phonetic reality (e.g. Dille, Ladd & Schepman, 2005). However, in the Gussenhoven-influenced analyses on which IViE is based, the ‘trailing’ tone is less strictly defined in terms of a fixed distance with the ‘starred’ tone of the pitch accent. Therefore, this paper is not a comparison of the ToBI and IViE systems but the differences between them in regard to their inclusion/omission of the phrase tone are useful to highlight.

Gussenhoven's analysis justifies the use of the 'trailing' tone on the basis of a number of issues, in particular his belief (1984: ch3) that the inclusion of the 'tail'¹ in the 'nuclear' tone in the older British School of Intonational analysis is correct and therefore that AM (Autosegmental-Metrical) analyses should incorporate this into their representation. Therefore, the Gussenhoven-influenced analyses do indeed presuppose that the 'starred' tone and 'trailing' tone are somehow linked together but precise phonetic realisation rules for this have not been clearly defined. Given that in this school of thought, the issue of a fixed distance has strangely remained unaddressed and in the absence of a clear alternative phonetic motivation, the fixed distance view of the 'trailing' tone will be adopted in this paper.

1.2. *Irish English Intonation*

The intonation of Irish English is an extremely under-researched area. In this regard, the IViE project was a pioneering work as it represented the first occasion upon which Irish English was analysed in a contemporary AM framework. The varieties studied were Belfast and Dublin (Malahide). Previous work on Irish English Intonation largely consists of isolated impressionistic observations and a small amount of work in the older British School of Intonational analysis (e.g. Jarman & Cruttenden, 1976). Despite the inroads that the IViE project made, however, the majority of Irish English varieties remain completely untouched as far as intonational analysis is concerned.

1.3. *Phonetic alignment: two hypotheses*

The Constant Interval Hypothesis posits that a rather fixed distance exists between relevant *F0* turning points (peaks and valleys in the intonation contour). Support was found for this hypothesis in Ashby's (1978) paper on (British) English, which included the study of nuclear falls (support was also found in Local & Pearson (1980), which makes reference to Belfast English). Ashby found that despite differing numbers of syllables after the nucleus, the main part of the fall was completed in pretty much the same amount of time on the nuclear syllable itself (1978: 334-5). Therefore, when a 'tail' was present, there was a rapid fall on the nuclear syllable, followed by a levelling out. More recent support comes from Frota (2002), as noted in Gussenhoven (forthcoming), who claimed to find support for this hypothesis in relation to the two tones in the accent she labelled as H*+L (although not for the two tones in the accent she labelled as H+L*).

However, more support recently has been found for the Segmental Anchoring Hypothesis, particularly in studies of rises (e.g. Arvaniti, Ladd & Mennen, 1998; Ladd *et al.*, 1999; Dille, Ladd & Schepman, 2005). This hypothesis posits that *F0* turning points are aligned with respect to specific segmental points (e.g. consonant/vowel points, syllable edges, etc.) rather than being aligned with respect to each other in time. In fact Arvaniti, Ladd and Mennen's study included an experiment specifically designed to test whether an edge-aligned phrase tone or fixed distance 'trailing' tone was appropriate for Greek prenuclear rises. On the whole though, the study of phonetic alignment in (nuclear) falls has received much less attention in the literature, so there is further scope for examining both hypotheses in this context.

2. Stage 1: Analysis of IViE Dublin Data

The data analysed in this section represents a subset of the Dublin data acquired in the IViE project (Grabe, Post & Nolan, 2001). There were twelve Dublin speakers in total and they

¹ The unaccented part of the intonation contour after the nuclear syllable (main and usually final accented syllable in the IP).

were recorded in Malahide, approximately seven miles north of Dublin city centre (Malahide English is quite a different variety from the English of central Dublin). The subset of data used in the present paper is a set of 87 declaratives taken from the set of read sentences. The declarative utterances used are those featuring a nuclear fall and they were analysed using PRAAT (Boersma & Weenink, 2005).

2.1. Is the fixed distance notion tenable?

In order to evaluate the appropriateness of the Constant Interval Hypothesis for Malahide nuclear falls, one method is to examine what happens when there are varying numbers of syllables between the nuclear syllable and the IP boundary. This method was used in Ashby's (1978) study of nuclear falls and it was also used in relation to rises in Arvaniti and Ladd's (1995) study of Greek. In 81 of the 87 utterances analysed in the present experiment, the nuclear accent was located on the final word of the IP, not unexpected given that in these read sentences, no special emphasis (e.g. contrastive focus) was elicited. In the other six utterances, the nuclear accent was clearly produced away from the boundary.

Taking the 81 utterances as a group and the 6 exceptions as a group, the distance between the relevant *F0* max (H) and min (L) was much greater in those cases where the nuclear accent is away from the boundary. If there was a fixed distance between these two points we might expect H to migrate further rightwards as the number of syllables between the accented syllable and the IP boundary increased, preserving the shape of a particular section of *F0* contour. Indeed one of the numerous factors cited in the phenomenon of rightward peak delay² has been the number of syllables following the nuclear syllable (Steele, 1986). However, peak delay did not occur in the six exceptional cases in this data. Neither did these utterances manifest the shape of the utterances in Ashby's data whereby a rapid fall occurs on the nuclear syllable itself and any following syllables are produced on a fairly level tone. The mean duration between H and L for 81 utterances (nuclear accent at boundary) was 263 ms (milliseconds) (SD 79). However, when the nuclear accent is away from the boundary the mean distance was 541 ms (SD 101) (note: in all cases where the nuclear accent was away from the boundary, there were 3 syllables between it and the IP boundary). The difference between the two means casts some initial tentative doubt on the Constant Interval Hypothesis. A one-tailed Wilcoxon-Mann-Whitney Test was carried out on this data, giving a *z* value of 3.99. This rejected the hypothesis that the two groups of sentences ('nuclear-final', 'nuclear-non-final') could have come from the same distribution, and confirmed that the distance between the *F0* max and min was greater in 'nuclear-non-final' group at a significance level $p < 0.0005$.

2.2. Malahide alignment

H was found to be aligned consistently closer to the beginning of the rhyme of the accented syllable (v1) than to the syllable onset (c1), suggesting some stability of alignment of the type predicted by those in support of the Segmental Anchoring Hypothesis (e.g. Arvaniti, Ladd & Mennen, 1998). A one-tailed Wilcoxon-Mann-Whitney test confirmed this (*z* value 2.74, $p < 0.005$). L was located consistently after the rhyme of the IP-final syllable (v2). A one-tailed Wilcoxon test showed that L was located closer to the end of the IP (E) than to v2 or some other specific segmental landmarks. Nevertheless, there was considerable variability both in the distance between v1 and H (mean 44 ms, SD 49) and between L and E (mean 57 ms, SD 51). This would not support the strictest version of the Segmental Anchoring Hypothesis.

² This occurs when the peak is located beyond the accented syllable.

2.3. Discussion

From these initial experiments, we have found some more support for the Segmental Anchoring Hypothesis than for the Constant Interval Hypothesis, although we did not find full support for either hypothesis. We need to examine now how this relates to our query of whether there is a reflex of a phrase tone or a 'trailing' tone at L. We do not believe that we have enough evidence at this stage to give unreserved support for the phrase tone analysis. The main reason for this is that the *F0* trace in the vast majority of these Malahide utterances exhibits an extremely gradual fall without phonetic evidence of two distinct pitch movements after the *F0* max (i.e. phonetic reflexes of both a phrase tone and a boundary tone). It should be added though, that a strict ToBI analysis of the data would posit both a phrase tone and a boundary tone at the IP boundary even when both cannot be phonetically detected. A possible alternative solution is to adopt the labelling used by Hayes and Lahiri (1991) for nuclear falls in Bengali: H* L%. That is, omitting the phrase tone and simply positing a low boundary tone after the high pitch accent. This seems more appropriate for the Malahide data and avoids the 'trailing' tone's association with the notion of a fixed distance. So at this stage, we believe that our analysis does not support the inclusion of either a phrase tone or a 'trailing' tone.

3. Stage 2: Analysis of Wexford English

This variety of English (Wexford is located in the southeast of the Republic of Ireland) has probably never been analysed intonationally before and certainly not in a contemporary AM framework. Data was collected from four female speakers, all undergraduate students at University College Dublin. Each speaker was recorded in a quiet room in this university. A Marantz digital recorder Model PMD670 and a Shure microphone Model SM10A were used. There were no reports of any speaker having speech or hearing difficulties.

3.1. Corpus design

Each speaker was recorded in a set of read sentences, designed specifically for this experiment (i.e. not the same sentences as used in the IViE project). There were thirty sentence units of three different types. Each speaker was recorded reading the complete list of sentences, twice, with the sentences in random orders. There were also three extra buffer sentences placed at the beginning and end of each sentence list, which were not analysed. The types of sentences recorded were:

1. simple declaratives,
2. complex declaratives (in which yes/no questions were embedded into a declarative structure) and
3. declaratives in the context of a question and answer (Q&A) pair.

Types 2 and 3 were designed specifically to elicit deaccenting³ of repeated material and thus by contrastive focus elicit the nuclear accent away from the IP boundary. A total of 240 sentences were recorded (4 speakers each reading a list of 30 sentences, twice). A few sentences had to be discarded, due primarily to production errors in reading, leaving 230 sentences available for analysis.

3.2. Simple declaratives

In all simple declaratives, the nuclear accent was elicited on the final word of the IP. The simple declaratives were devised with an alternating number of unstressed syllables (0 to 2) following the accented syllable. This allows for an initial examination of whether the fixed distance between the *F0* max and min is preserved with the increase in syllables after the

³ This phenomenon also occurs in French. See Post, 2000: 151 ch. 6.

accented one. One-tailed Wilcoxon-Mann-Whitney tests were carried out for each speaker and show some evidence of an increase in distance between H and L alongside the increase in syllables: RF (z value 3.344, $p < 0.0005$), GH (z value 2.855, $p < 0.005$), LM (z value 2.389, $p < 0.01$), NR (z value 1.911, $p < 0.05$). This suggests that the notion of a fixed distance between the *F0* max and min in this Wexford data is dubious.

3.3. *Wexford alignment: F0 maximum*

H was found to be aligned closer to c1 than v1 in the simple declaratives in three out of the four speakers. This could possibly represent a dialect difference with Malahide as speakers aligned H closer to v1 there. This must be very tentatively stated, as the goal of this present paper was not to capture the difference between these two varieties. However, only two speakers showed evidence of aligning H significantly closer to c1 than v1 (GH z value 2.1, $p < 0.01$, LM z value 1.8, $p < 0.05$). In addition, there was considerable variability in the distance between c1 and H in these two speakers: GH (Mean 28 ms, SD 21); LM (Mean 27 ms, SD 28). In the complex declaratives and Q&A pairs, there was rightward *F0* peak 'migration'⁴. It is most likely that the degree of emphasis expressed on the accented word due to the elicitation of contrastive focus, is what resulted in this phenomenon here (cf. Gussenhoven, 1999; Ladd & Morton, 1997). In addition, the greater number of syllables between the accented word and the IP boundary (Steele, 1986) in the complex declaratives/Q&A pairs may have had some effect.

3.4. *Contour shape in Q&A pairs*

An interesting contour shape emerged in the Complex Declaratives and Q&A Pairs in three speakers (NR, RF, LM). A rapid fall from the *F0* peak was followed by a 'plateau' (much more gradual fall) which lasted until the end of the IP. This contour shape must be accounted for. Referring to the analysis of similar contours by others, it is possible that the turning point at the end of the rapid fall could represent the secondary association of an L phrase tone, whose main docking site is nearer the IP boundary (Pierrehumbert & Beckman, 1988).

The location of the turning point at the end of the rapid fall (SL) was of particular interest. For all three speakers who displayed the rapid fall-'plateau' pattern, the turning point (SL) seemed quite consistently located towards the end of the first syllable after the end of the accented word. For RF, 76% of relevant utterances had SL located in this place; for LM, the figure was 72% and for NR, 66%. Considering the relevant utterances of all three speakers together, the figure is 71%. This is unusual in that it does not correspond strictly to Pierrehumbert's (1980) suggestion that the phrase tone would be aligned at the edge of the accented word. The data studied also did not appear to show an effect of the first postnuclear stressed syllable, which was found to be the docking site of the L phrase tone in Grice, Ladd and Arvaniti (2000). The 2000 paper also supported the view that there would be secondary association of the L phrase tone at the edge of the accented word. Measurements were taken to examine if there was a fixed distance between the *F0* peak (H) and the turning point at the end of the rapid fall (SL). One-tailed Wilcoxon-Mann-Whitney tests were carried out for each speaker comparing the distance between H and SL where the accented word contained one and two syllables respectively. If the distance was not fixed, we might expect the distance between H and SL in the disyllabic accented words to be greater. However, we did not obtain significant results for any speaker (LM, z value 0.127; RF, z value 1.239; NR, z value 0.021). This means that we cannot rule out the possibility of a somewhat fixed distance between these points. Yet all speakers displayed a considerable degree of variation in the

⁴ We use this term 'migration' to distinguish what we noticed in our data from the phenomenon of 'peak delay' as the peak did not occur after the accented syllable in our data.

distance between H and SL, which would suggest that the distance could not be absolutely fixed (RF mean distance H to SL 180 ms (SD 65); LM 187 ms (SD 89); NR 261 ms (SD 146)). In addition, we must attempt to reconcile the possibility of a fixed distance between H and SL with the fact that SL seemed to have a fairly consistent location and thus may be segmentally anchored. Clearly further work is needed. What is interesting though, is that however we define the relationship between H and SL, the location of SL does not correspond to any kind of prosodic boundary. The implication of this is that it makes it difficult to conceptualise SL as a phrase tone.

3.5. *Wexford alignment: F0 minimum*

The alignment of the true *F0* min (final L) was also examined. In the majority of utterances from each speaker, L was located after v2 yet before the final coda consonant (c3). Kruskal-Wallis One-Way ANOVAs by ranks were carried out on the data of each speaker to see if L was favouring one of a number of segmental points. The results were significant for three speakers (RF, χ^2 40.126, $p < 0.001$; NR, χ^2 24.93, $p < 0.001$; GH χ^2 12.338, $p < 0.01$). Yet we are stilling dealing with a high degree of variability (mean distance from v2 to final L: all speakers together 84 ms (SD 73), RF 86 ms (SD 69), GH 85 ms (SD 72), LM 65 ms (SD 50) and NR 98 ms (SD 89). What is interesting about the alignment of this final L is that it is not certain that it is aligned with respect to the end of the IP (or another prosodic boundary), as for two speakers at least the locations of L tended to cluster around v2. This raises the possibility that it is not behaving as though it were an IP boundary tone (L%) (i.e. a delimiter of the IP). Grønnum (1992) in her excellent critique of van den Berg, Gussenhoven and Rietveld's chapter on Downstep in Dutch, argues for getting rid of L%, suggesting that the slight final falls to which she refers, "be ascribed to weak edge vibrations before a pause" (1992: 364). On this point, IViE's optional specification of boundary tones, allowing for a boundary to be specified 0% where no significant change in contour direction takes place, seems entirely appropriate in the case of the Wexford data (and perhaps also the Dublin data noted earlier). The 0% option thus takes into consideration Grønnum's observation, and it also avoids the problem of applying a boundary tone label to a tone which may not be thus aligned, though the phrase tone label (as with SL above, with its implication of a prosodic boundary) still remains troublesome.

4. Conclusions

The data studied from Dublin (Malahide) and Wexford English have been shown to give more support overall for the Segmental Anchoring Hypothesis than the Constant Interval Hypothesis. This cannot be translated so easily though, into giving automatic support for an analysis involving phrase tones over one involving 'trailing' tones. A problem for the ToBI analysis in relation to phrase tones (and boundary tones) is that neither the true *F0* min (final L) in the Wexford data, nor the elbow at the end of the rapid fall (SL) in the Wexford data, appear to function as edge tones to any kind of prosodic phrases. The ToBI system explicitly states the phrase tone and boundary tone occur at the end of the intermediate and intonational phrases respectively. Thus, we have the ambiguous situation that full support has not been found for either the phrase tone or the 'trailing' tone analysis in this Irish English data. For the Dublin (Malahide) data we suggested omitting the phrase tone altogether and considering the *F0* min a low boundary tone (L%). However, this solution does not appear to be appropriate for the Wexford data. We could not rule out the possibility of some kind of fixed distance between H and SL in Wexford, so a 'trailing' tone analysis might be more appropriate here. We would be reluctant to imply the rigid fixed distance that is incorporated in to the ToBI notion of a 'trailing' tone and would like to explore the somewhat different concept of the 'trailing' tone in Gussenhoven's (1984) system. The notion of the secondary

association of the phrase tone, which was raised earlier as a possibility for the Wexford data is difficult to support for two main reasons:

- it was uncertain that L in Wexford was behaving as a phrase tone in the first place;
- the location of SL did not correspond to the edge of the accented word, which was the proposed location for the secondary association.

Our original question was whether the L tone in nuclear falls in these two dialects of Irish English would be better represented as phrase tones or 'trailing' tones. In fact, our research has led us away from phrase tone labels. This is somewhat ironic as overall we found more support (though not full support) for the Segmental Anchoring Hypothesis in our data and other authors have used this to argue for a phrase tone analysis (e.g. Grice, Arvaniti & Ladd, 2000). It is very interesting to put our findings in the context of systems (e.g. IViE system) which avoid phrase tones altogether. We believe that this work highlights the need for a better understanding of the relationship between the phonetic details of the intonation contour and phonological categorisation. As noted by other authors (e.g. Arvaniti, Ladd & Mennen, 2000; Dilley, Ladd & Schepman, 2005) the correspondence between tonal labels and turning points in the contour cannot be as direct as once thought. Our research, though still at an early stage, would suggest that phrase tones may not be needed to represent Dublin (Malahide) and Wexford nuclear falls. We believe that this point is significant in itself, even though ambiguities have remained about the precise characterisation of the 'trailing tone'.

Bibliography

- ARVANITI, Amalia, LADD, D. Robert & MENNEN, Ineke (1998). Stability of tonal alignment: the case of Greek prenuclear accents, *Journal of Phonetics*, 26 : 3-25.
- ASHBY, Michael (1978). A study of two English nuclear tones, *Language and Speech*, 21 : 326-336.
- BECKMAN, Mary E. & PIERREHUMBERT, Janet (1986). Intonational structure in Japanese and English, *Phonology Yearbook*, 3 : 255-309.
- BRUCE, Gösta (1977). *Swedish word accents in sentence perspective*. Lund : CWK Gleerup.
- DILLEY, Laura C., LADD, D. Robert & SCHEPMAN, Astrid (2005). Alignment of L and H in bitonal pitch accents: testing two hypotheses, *Journal of Phonetics*, 33 : 115-119.
- GRICE, Martine, LADD, D. Robert & ARVANITI, Amalia (2000). On the place of "phrase accents" in intonational phonology, *Phonology*, 17 : 143-185.
- GRØNNUM, Nina (1992). Comments on R. van den Berg, C. Gussenhoven & T. Rietveld, 'Downstep in Dutch: implications for a model', in *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*, Docherty, Gerald J. & Ladd, D. Robert (Eds). Cambridge: CUP, 359-367.
- GUSSENHOVEN, Carlos (1984). *On the grammar and semantics of sentence accents*. Dordrecht : Foris.
- GUSSENHOVEN, Carlos (1999). Discreteness and gradience in intonational contrasts, *Language and Speech*, 42 : 283-305.
- GUSSENHOVEN, Carlos (Forthcoming). The Phonology of Intonation, in *The Cambridge Handbook of Phonology*, De Lacy, Paul (Ed). Cambridge: CUP, 399-444.
- HAYES, Bruce & LAHIRI, Adita (1991). Bengali intonational phonology, *Natural Language and Linguistic Theory*, 9 : 47-96.
- JARMAN, E. & CRUTTENDEN, Alan (1976). Belfast intonation and the myth of the fall, *Journal of the International Phonetic Association*, 6 : 4-12.
- LADD, Robert D., FAULKNER, Dan, FAULKNER, Hanneke & SCHEPMAN, Astrid (1999). Constant segmental anchoring of *F0* movements under changes in speech rate, *Journal of the Acoustical Society of America*, 106 (3): 1543-1554.
- LADD, Robert D. & MORTON, Rachel (1997). The perception of intonational emphasis: continuous or categorical?, *Journal of Phonetics*, 25 : 313-342.

- LICKLEY, Robin J., SCHEPMAN, Astrid & LADD, D. Robert (2005). Alignment of “Phrase Accent” lows in Dutch falling rising questions: theoretical and methodological implications, *Language and Speech*, 48 (2): 157-183.
- LOCAL, John & PEARSON, Mark (1980). Some physical properties of three English tones, *York Papers in Linguistics*, 8 : 207-218.
- PIERREHUMBERT, Janet B. (1980). *The phonetics and phonology of English intonation*. PhD Thesis, MIT.
- PIERREHUMBERT, Janet B. & BECKMAN, Mary E. (1988). *Japanese tone structure*. Cambridge, MA: MIT Press.
- POST, Brechtje (2000). *Tonal and phrasal structures in French intonation*. The Hague, Holland Academic Graphics.
- SILVERMAN, Kim, BECKMAN, Mary, PITRELLI, John, OSTENDORF, Mari, WIGHTMAN, Colin, PRICE, Patti, PIERREHUMBERT, Janet & HIRSCHBERG, Julia (1992). ToBI: A standard for labelling English prosody, *Proceedings of the 1992 International Conference on Spoken Language Processing* 12-16 October, Banff, Canada: 867-870.
- STEELE, Shirley (1986). Nuclear accent F0 peak location: Effects of rate, vowel, and number of following syllables, *Journal of the Acoustical Society of America*, 80 (Suppl. 1) : S51.

References to internet sites

- GRABE, Esther, NOLAN, Francis & POST, Brechtje. English Intonation in the British Isles, <http://www.phon.ox.ac.uk/IViE/>, (2003). Consulted in March 2007
- GRABE, Esther, POST, Brechtje & NOLAN, Francis. The IViE Corpus. Department of Linguistics, University of Cambridge, <http://www.phon.ox.ac.uk/IViE/>, (2001). Consulted in March 2007

References to software

- BOERSMA, Paul & WEENINK, David (2005). *Praat: A system for doing phonetics by computer (Version 4.3.22) [Computer program]*. <http://www.praat.org>. Retrieved September 8, 2005.

LE PRÉSENT EXPRIME-T-IL L'ACTUALITÉ ?

LA DOUBLE PERSPECTIVE DE L'ENCODAGE ET DU DÉCODAGE

Christian SURCOUF

christian.surcouff@u-grenoble3.fr

Laboratoire LIDILEM

Université Stendhal, Grenoble 3

Abstract : This article challenges the claim that the French present tense lacks temporal value. Though it also encodes information, reporting a fact that is contemporaneous with the moment of speech *always* requires the present tense in elementary speech situations, i.e. "Nowness" entails present tense. Therefore a notion of "nowness" pervades almost all uses of this tense, mainly *de dicto* and occasionally *de re*.

1. Introduction

Dans son article intitulé *Le prétendu "présent" de l'indicatif : une forme non déictique du verbe*, Serbat (1988 : 33) déclare : « Le présent est rigoureusement étranger à la notion d'actuel, et, en général à toute notion d'époque ». En s'étayant notamment sur la constatation de Kahn (1954 : 59) : « Dans les 2000 phrases, nous n'avons trouvé que 7 exemples de la simultanéité à peu près stricte du présent », Touratier (1996 : 95) conclut également : « un verbe au présent n'a par lui-même aucune valeur temporelle et ne contient nullement le morphème de signifié 'actuel' que les grammaires scolaires voudraient lui attribuer ».

Nous essaierons de montrer que ces affirmations résultent d'un raisonnement fusionnant deux perspectives divergentes et qu'en définitive le Présent¹ renvoie en son essence à l'actuel.

Pour définir l'assise temporelle sur laquelle se construit la communication verbale, nous invoquerons en premier lieu la notion de *présent psychologique* à l'origine de l'ancrage spatiotemporel caractéristique du *moment d'énonciation*. Dans le prolongement de ces considérations, la prise en compte de la situation pragmatique de communication permettra d'aborder la problématique du *temps grammatical* à travers le Présent, envisagé selon les perspectives de l'encodage et du décodage² de l'information dans une optique résolument référentialiste. En effet, la perception de la temporalité ne paraît pas concevable en dehors de la réalité à laquelle elle réfère.

2. Le présent psychologique et le moment d'énonciation

Malgré la place privilégiée qu'il occupe dans le fonctionnement des temps grammaticaux, le moment d'énonciation ne paraît guère susciter l'intérêt des linguistes. En témoigne notamment l'absence de consensus sur sa nature, oscillant entre *point* (p. ex. Reichenbach, 1947) et *intervalle* (p. ex. Gosselin, 1996). Une première caractéristique semble pourtant s'imposer : il s'inscrit toujours dans le *présent du locuteur*, notion extralinguistique que la psychologie expérimentale va nous aider à mieux circonscrire. Notre conscience du présent, *présent psychologique* (Fraisie, 1957/1967) ou *présémantique* dans la terminologie de Pöppel s'assimilerait ainsi à une : « temporal platform for conscious activity [...] of

¹ Écrit avec une majuscule, « Présent » renvoie au tiroir verbal.

² Dépourvue de biunivocité, la langue n'est pas un code. On utilisera toutefois les termes *encodage* et *décodage* pour leur symétrie.

approximately 2 to 3 seconds [...]. This universal integration process is automatic and pre-semantic, i.e. it is not determined by what is processed [...]. [It] provides a logistic basis for conscious representation, a working platform for our phenomenal present » (Pöppel, 2004: 298 & 300). Le *sens* permettrait alors d'assurer la continuité entre ces intervalles : « continuity is assumed to be the consequence of a semantic connection of what is represented within each subjective present » (Pöppel, 1997 : 60).

D'un point de vue linguistique, on peut tirer plusieurs conséquences de l'existence du présent psychologique : (1) il permet à toute énonciation d'être ancrée dans notre conscience *universelle* de l'espace-temps et (2) implique par là même une perception *partagée* de cet ancrage ; (3) l'évocation d'une durée supérieure à 3 secondes transiterait nécessairement par le sémantisme. Ces trois points établis, avant de mener notre réflexion sur le Présent (temps grammatical), définissons la situation de base de toute communication.

3. La situation de discours pur

D'un point de vue ontogénétique (et vraisemblablement phylogénétique), la langue ne semble pouvoir émerger qu'à partir de la situation minimale mettant *face à face* deux êtres humains s'exprimant oralement. Le locuteur et l'interlocuteur embrassent ainsi un même espace-temps étayé sur une conscience partagée – parce qu'universelle – du présent psychologique. Cette configuration par défaut constitue ce que nous appellerons la « situation de discours pur » (des réflexions dans ce sens existent chez Adam, Lugin & Revaz, 1998). Quels que soient les apports technologiques permettant le décalage spatiotemporel entre communicants (enregistrements scripturaux, magnétiques, informatiques, etc.), en tant que phénomène linguistique, le temps grammatical ne peut avoir émergé qu'à partir de cette configuration par défaut, base de notre réflexion.

4. Lexis et durée

Examiner l'expression de la temporalité *par* la langue requiert la prise en compte d'au moins deux phénomènes : (1) l'ancrage dans l'espace-temps du fait³ relaté, ainsi que (2) sa durée. Occupons-nous pour l'instant du second en précisant tout d'abord qu'il ne s'agit pas d'une capacité inhérente à la langue d'évoquer *en soi* une quelconque durée. Par exemple, « étincelle » n'est pas plus long ou plus court que « flamme ». Seule notre représentation du monde nous autorise à penser que le référent prototypique du premier évoque une brièveté, improbable chez le second. Comme les recherches sur l'aspect lexical l'ont mis en évidence (voir notamment l'article fondateur de Vendler, 1957, et les développements qu'il a suscités), les verbes ou les syntagmes verbaux eux aussi peuvent donner lieu à un comportement comparable. Ainsi « apercevoir » (classé parmi les « achèvements ») renverrait à une action plus fugace que « regarder » (« activité »)⁴. À une échelle supérieure mais toujours en amont de l'énoncé, ce constat vaut également pour la lexis⁵ (signalée entre crochets <>) : <La toupie/tourner> renvoie à un évènement plus court que <La Terre/tourner>. Signalons enfin que même considérée en dehors de toute utilisation concrète, quel que soit le contexte, une lexis comme <Léo/apercevoir sa mère dans le jardin> décrira toujours un procès plus court

³ « Fait » est ici pris dans un sens large comme « ce qui peut être exprimé au moyen de la langue ».

⁴ Même si elle comprend souvent le trait [\pm ponctuel] ou son pendant [\pm duratif], la catégorisation en modes de procès ne procède pas uniquement de l'inférence de la durée (voir p. ex. Vetters (1996 : 87-106) ou Fuchs (1991 : 9-17)).

⁵ Dans le sens de Culioli (1990 : 78) : « A lexis is not an utterance (énoncé). It is neither asserted nor unasserted, for it has not yet been situated (or located) within an enunciative space ».

que la durée nécessaire à l'articulation de l'ensemble des mots servant à l'exprimer⁶. Bien que partiel, cet aperçu montre qu'une orientation temporelle, en l'occurrence la durée, existe déjà au niveau du mot, du syntagme et de la lexis (l'importance de ce constat apparaîtra dans les exemples 5.1.1 à 5.1.4 en 5.1). D'autre part, en ce qui concerne (1), comme le remarque Klein (1994 : 36), « Lexical contents by themselves are not directly linked to time. [...] the content <*the apple fall on the head*> is not earlier or later than the content of <*Peter eat a banana*> ». Pour pouvoir renvoyer à une période précise, la lexis doit être repérée dans le temps. En l'absence de chronologie universelle absolue qui rendrait possible un tel repérage, seul le moment d'énonciation – partagé grâce à la conscience intersubjective du présent psychologique – pourra, en situation de discours pur, remplir cette fonction.

Abordons maintenant la problématique du Présent qui, avec les déictiques temporels, permet de situer le « fait » relaté par rapport au moment d'énonciation.

5. L'effectivité du sémantisme de la lexis et le Présent

5.1. La perspective de l'encodage

Envisageons une situation de discours pur (cf. 3) où A, témoin direct de la scène suivante, désire en faire part verbalement à B :



Schématiquement, on peut imaginer les étapes ci-dessous :

1. A identifie Léo et construit la lexis <Léo/cuisiner>.
2. A transforme cette lexis en phrase pour signifier à B que son contenu sémantique renvoie à une réalité contemporaine du moment d'énonciation.
3. A énonce : « Léo cuisine. ».

Une remarque importante s'impose ici. À l'exception du Présent (ou de sa forme progressive en « être en train de »), A ne peut recourir à aucun autre tiroir verbal. En effet, parmi les énoncés suivants, aucun ne permet la description de cette réalité : Léo cuisinait/avait cuisiné/a cuisiné/cuisina/eut cuisiné/cuisinera/aura cuisiné/va cuisiner/vient de cuisiner/aurait cuisiné/cuisinerait, etc. Dans la situation de discours pur, en *encodage*, s'ensuit donc l'implication suivante :

1 Le sémantisme de la lexis renvoie à une réalité contemporaine du moment d'énonciation	⇒	Présent
---	---	---------

Comme en témoignent les exemples 1 à 4, ce principe demeure valide en dépit de la grande disparité de durée impliquée par les lexis :

1. <Je/te baptiser Aélia> ⇒ Je te baptise Aélia.
2. <Je/avoir un frisson> ⇒ J'ai un frisson.
3. <Je/avoir les yeux verts> ⇒ J'ai les yeux verts.
4. <La Terre/tourner> ⇒ La Terre tourne.

Si les durées de 1 et 2 peuvent se concevoir comme directement perceptibles grâce à notre conscience du présent, il est évident que même en y adjoignant la mémoire, il est impossible de percevoir *en tant qu'unité* l'intégralité des instants sous-entendus par 3 et 4. Seul le

⁶ Cet « inconvénient » inhérent à tous les achèvements rend délicate leur narration *en direct*. Par exemple, dans : « Il entre dans sa voiture, claque la portière et démarre. », il est probable que le récit soit en retard (ou en avance par anticipation déterministe de la part du locuteur) sur le déroulement réel des événements. En ce sens l'utilisation du Présent pour les achèvements s'avère toujours problématique puisque l'événement est susceptible de se terminer avant l'énoncé qui le relate. Ceci expliquerait qu'un énoncé tel que « J'arrive à l'instant. » soit synonyme de « Je viens d'arriver. », alors que « Je cours à l'instant. » est en revanche impossible. Pour cette raison, notre raisonnement s'étaiera essentiellement sur les états, accomplissements et activités (cf. §7).

contenu sémantique de la lexis basé sur notre connaissance du monde permet d'y parvenir. Aussi adhérons-nous au point de vue de Comrie : « the present tense refers only to a situation holding at the present moment, even when that situation is part of a larger situation that occupies more than just the present moment ». En effet, en 3 et 4, le Présent se contente de signaler que le sémantisme des lexis <Je/avoir les yeux verts> et <La Terre/tourner> renvoie à une réalité *contemporaine* du moment d'énonciation (pour un raisonnement similaire à propos de « now », voir Origgi & Sperber, 2000 : 158). En conséquence, il est normal que, parmi les 2000 phrases de son corpus, Kahn (1954 : 59-60) ne trouve que sept cas de « simultanéité plus ou moins stricte entre la durée du moment de la parole et la durée référée » (ici, elle serait stricte avec le performatif en 1 et, probable en 2). Par ailleurs, contrairement aux affirmations de Serbat (1988) et de Touratier (1996), en encodage, comme l'attestent ces exemples, le Présent *n'est pas* « étranger à la notion d'actuel » (Serbat, 1988 : 33) ; c'est en fait le seul tiroir verbal à permettre d'explicitier l'actualité d'une information⁷. En somme, les arguments avancés par Kahn, Serbat et Touratier pour nier au Présent sa valeur d'actualité semblent faire abstraction de cette première donnée et puiser leur justification dans la perspective du décodage.

5.2. La perspective du décodage

Si en encodage, lorsque le sémantisme de la lexis renvoie à une réalité contemporaine du moment de l'énonciation, l'énoncé appelle obligatoirement l'utilisation du Présent (« Léo cuisine. »), lors du *décodage*, la réciproque n'est pas toujours vraie. En témoignent notamment les deux contre-exemples suivants :

5. Demain, Léo cuisine.

6. Hier, je rentre chez moi, Léo cuisine un petit plat⁸, et puis tout à coup il me dit...

En 5 et 6, malgré l'emploi du Présent dans l'énoncé, l'interprétation qui procéderait du cheminement inverse de l'implication **1** est bloquée par l'utilisation des déictiques temporels *demain* et *hier*. Faisant office d'« introducteurs » à un « espace mental » (voir Fauconnier, 1984 : 32-33), *hier* et *demain* projettent l'interlocuteur dans un univers non-contemporain du moment d'énonciation et neutralisent ainsi le sémantisme du Présent observé dans l'implication **1**, que nous rappelons ici :

1 Le sémantisme de la lexis renvoie à une réalité contemporaine du moment d'énonciation ⇔ Présent

En ce sens, en décodage, étant donné les restrictions relevées dans les exemples 5 et 6, l'implication réciproque de **1** n'est pas valide :

2 Présent ✗ Le sémantisme de la lexis renvoie à une réalité contemporaine du moment d'énonciation

En somme, l'équivalence suivante est invalide :

1 & 2 Présent ✗ Le sémantisme de la lexis renvoie à une réalité contemporaine du moment d'énonciation

Cependant, comme l'atteste l'implication **1**, il n'est pas possible d'en déduire, à l'instar de Serbat (1988 : 33) que « le présent est rigoureusement étranger à la notion d'actuel, et, en général à toute notion d'époque ». En situation de discours pur, relater un état du monde

⁷ Les quelques rares exceptions demanderaient une analyse approfondie qui sort du cadre de cet article. Propres à des situations bien spécifiques, signalons par exemple les énoncés : « Je voulais/voudrais/aurais voulu un pain. », « J'avais une question ! ». On peut d'ailleurs d'emblée remarquer que : « Il voulait un pain » ne semble déjà plus interprétable comme actuel.

⁸ Apparaît ici une première restriction dans l'usage du Présent pour relater une histoire dans le passé. En effet, il est impossible de conserver l'*activité* <Léo/cuisiner> qu'il nous a fallu remplacer par l'*accomplissement* <Léo/cuisiner un petit plat>. Le maintien de l'activité aurait demandé le recours à l'Imparfait : « Hier, je rentre chez moi, Léo *cuisinait*, [...] » (à ce propos, voir CULIOLI 1999 :136). De plus, l'utilisation du Présent avec *hier* exige un marquage intonatif particulier et l'élaboration d'un récit composé de plusieurs propositions (cf. §6.1).

actuel, c'est-à-dire contemporain du moment d'énonciation requiert toujours le Présent (voir la note 7 pour quelques exceptions). Le Présent n'est donc pas « étranger à la notion d'actuel », il est au contraire le seul à autoriser son expression en encodage. Une position radicalement contraire à celle de Serbat s'avère tout aussi inappropriée : « Chaque fois qu'un locuteur emploie la forme grammaticale de "présent" (ou son équivalent), il situe l'évènement comme contemporain de l'instance du discours qui le mentionne. » (Benveniste, 1974 : 73). Ici, Benveniste semble considérer que l'implication **1** admettrait *toujours* (« chaque fois ») une réciproque, qui aboutirait alors à la validité de l'équivalence **1 & 2**. Bien qu'un énoncé tel que « Léo cuisine. » puisse effectivement impliquer la contemporanéité de l'activité culinaire de Léo avec le moment d'énonciation, comme les exemples 5 et 6 le démontrent, l'utilisation du Présent ne permet pas *systématiquement* de situer « l'évènement comme contemporain de l'instance du discours qui le mentionne ». La reformulation suivante de la citation de Benveniste rectifierait la confusion et correspondrait alors à l'implication **1** : « Chaque fois que le locuteur veut exprimer un évènement contemporain de l'instance du discours, il *doit* employer le Présent ».

6. Incidence des indications temporelles sur le Présent

6.1. Le cas du déictique temporel passé : hier

Si, comme le soutient Touratier (1996 : 96), le Présent est « dépourvu de signifié propre [et] est [...] apte à apparaître dans les types d'énoncés les plus divers, voire les plus curieux au point de vue de la temporalité », comment expliquer les différences d'acceptabilité des deux exemples suivants ?

5. Demain, Léo cuisine.
7. *Hier, Léo cuisine.

En effet, contrairement à 5, 7 ne peut fonctionner de façon autonome. Quelle est donc l'origine de l'incompatibilité du Présent avec *hier* en 7 et de sa compatibilité avec *demain* en 5 ? Un changement dans la catégorie du mode de procès ne modifie en rien l'incompatibilité du Présent avec *hier* dans un énoncé autonome :

8. *Hier, Léo perd sa clé. (achèvement)
9. *Hier, Léo est malade. (état)
10. *Hier, Léo cuisine un petit plat. (accomplissement)

Si contrairement à Touratier, et conformément à l'implication **1**, on reconnaît au Présent une valeur d'actualité, il devient alors possible d'expliquer le conflit entre le Présent et *hier* dans les exemples 7 à 10. Examinons le cas de 7. En situation de discours pur, le déictique *hier* renvoie obligatoirement au passé et crée un espace mental devant inclure l'évènement évoqué par la lexis <Léo/cuisiner>. Or, la proposition au Présent « Léo cuisine. » renvoie par défaut à un évènement contemporain du moment d'énonciation, la rendant incompatible avec *hier*. En tant que déictique du passé, *hier* requerrait l'utilisation du Passé composé : « Hier, Léo a cuisiné ». Toutefois, comme l'attestait déjà l'exemple 6 (cf. note 8), la valeur d'actualité du Présent peut être contrecarrée par une forme de « mise en récit »⁹. Ce récit dramatisé¹⁰ devra comprendre un minimum de deux propositions au sein desquelles les verbes d'accomplissement et d'achèvement pourront apparaître au Présent. L'énoncé 12 se démarque ainsi de 11 par son acceptabilité.

11. *Hier, je rentre chez moi.
12. Hier, je rentre chez moi ; l'appartement était/*est complètement inondé !

⁹ En effet, ni le *temps de re* ni le *temps de dicto* ne sont ici actualisés (pour ces notions, cf. §6.2).

¹⁰ Par exemple, la suite « Hier, je rentre chez moi. Je ferme la porte. » ne semble pas répondre aux critères de « mise en intrigue » constitutifs du récit (à ce sujet, voir ADAM, 1992 : 51-58).

Penchons-nous maintenant sur la question du futur.

6.2. *Le cas du déictique temporel futur : demain*

Si l'on maintient que le Présent possède une valeur d'actualité, pourquoi s'utilise-t-il apparemment sans problème avec *demain* ?

Demain, Léo cuisine. (activité)

5. Demain, Léo rentre tard. (achèvement)

13. Demain, Léo est libre. (état)

14. Demain, Léo cuisine un petit plat. (accomplissement)

À l'inverse des occurrences avec *hier* (appelant le plus souvent l'utilisation du Passé composé ou de l'Imparfait) qui inscrivent le fait ou l'évènement dans l'espace-temps, l'espace mental introduit par *demain* relève toujours du conjectural et n'autorise pas cette inscription spatiotemporelle. Ainsi, alors qu'une assertion renvoyant au passé se révèle automatiquement¹¹ soit vraie, soit fausse (« Hier, Léo a cuisiné. » est vrai si l'évènement dénoté par la lexis <Léo/cuisiner> a eu lieu la veille du moment d'énonciation), dès qu'il s'agit de l'avenir, cette possibilité est exclue. L'énoncé « Demain, Léo cuisine. » ne peut faire l'objet d'aucun jugement de vérité ; il se contente en effet de stipuler *au moment d'énonciation* un ordre des choses présumé *par le locuteur* en vertu de ses connaissances du monde. Aussi, quel que soit le contexte, l'assertion suivante paraît-elle difficilement concevable en raison de l'imprédictibilité de l'évènement décrit :

15. ?? Demain, Léo a un accident de train.

Mais alors quel rôle joue le Présent dans les exemples 5, 0, 13 et 14 ? Peut-on encore parler de valeur d'actualité ? Pour répondre à cette question, il nous faut recourir aux notions de « temps de *dicto* » et « temps de *re* » reconsidérées par Martin. Dans le cas de l'utilisation du Présent seul (donc en conformité avec l'implication **1**), Martin (1985 : 25) explique :

Je tiens la proposition Pierre est à Paris pour vraie aussi longtemps (temps de dicto) que j'ai toutes raisons de penser que Pierre est effectivement à Paris [=temps de re]. [...] En d'autres termes, j'assume (p) aussi longtemps que sa vraisemblance ne me paraîtra pas en cause. Mais dès lors que Pierre aura quitté Paris, cette proposition sera fausse, et il me sera impossible de la prendre en charge. Temps de re et temps de dicto paraissent donc se confondre.

Si l'on transforme maintenant l'énoncé de Martin en « Demain, Pierre est à Paris. », alors le *temps de re* n'a plus d'existence puisque l'état de fait dénoté par la lexis <Pierre/être à Paris> ne s'inscrit pas dans l'espace-temps au moment de l'énonciation. Demeure néanmoins le *temps de dicto* durant lequel le locuteur tiendra pour vraie la proposition « Demain, Pierre est à Paris. ». En somme, en accord avec Le Goffic & Lab (2000 : 96) : « Le présent du p.p.f. [=praesens pro futuro] est celui de l'assertion : le sujet énonciateur affirme une vérité *actuelle* pour lui, bien que de réalisation (d'effectuation) différée » (nous soulignons). Ici encore c'est en vertu de sa valeur d'actualité que le Présent peut remplir cette fonction.

6.3. *Autres interférences sur le Présent*

Outre ces interférences engendrées par des indications temporelles explicites, certaines autres caractéristiques de la lexis peuvent également inhiber le *temps de re* au profit du *temps de dicto*. Considérons l'énoncé suivant :

16. Le midi, Léo cuisine.

¹¹ La vérité résulte toujours d'un consensus plus ou moins étendu à propos des connaissances du monde. Les connaissances du futur n'existent pas mais résultent de spéculations déterministes.

Bien que, comme l'atteste la proposition « Léo cuisine. », la lexis <Léo/cuisiner> ne puisse normalement référer qu'à un évènement unique, la spécification temporelle « le midi » lui confère d'emblée un statut multiple (comparable au phénomène de pluriel linguistique). Quoiqu'il soit effectivement possible de se prononcer sur la valeur de vérité de l'énoncé 16, déclarer qu'il est vrai n'implique pas que l'évènement dénoté par la lexis est contemporain du moment d'énonciation. En effet, seule la valeur de vérité de l'assertion au moment d'énonciation peut être évaluée. Comme dans les cas entrevus en 6.2, la composante actuelle du Présent renvoie ici au *temps de dicto*. En somme, plus on s'éloigne de la perspective référentialiste, plus on s'achemine vers l'abstraction, où seul le *temps de dicto* s'avère déterminant. Ainsi dans la définition proposée en 17, la notion même de *temps de re* n'a plus lieu d'être.

17. La Terre est une planète du système solaire.

7. Conclusion

Loin d'être exhaustif, cet examen montre que le Présent conserve dans la plupart des cas sa valeur d'actualité¹² issue de l'implication 1 caractéristique de l'encodage en situation de discours pur. Lui dénier cette valeur sous prétexte qu'il se rencontre dans une grande diversité de contextes temporels paraît dès lors peu justifié. S'il en était dépourvu, pourquoi engendrerait-il des conflits avec certains déictiques temporels comme dans : « *Hier, Léo cuisine. » ? Chevalier (1996 : 85) n'a-t-il pas raison de nous avertir qu'il vaut mieux :

évite[r] d'en arriver à cette posture pour le moins étrange [...] qui consiste, au vu de phrases pour partie semblables, pour partie distinctes, à imputer les différences [...] non point aux éléments effectivement dissemblables mais, au contraire, aux éléments communs à ces phrases.

Aussi semble-t-il plus raisonnable d'aborder la description des tiroirs verbaux en souscrivant à la remarque de Comrie (1985 : 39) : « it is misleading to use collocation with time adverbials as a mechanical test for establishing the meaning of tenses ». La détermination des traits caractéristiques du Présent devrait donc se baser sur l'observation prioritaire d'énoncés simples du type : « Léo crie. », « 2 et 2 font 4. », etc. Néanmoins, un énoncé dépourvu d'indications temporelles explicites ne garantit pas l'absence d'influence sur son comportement temporel. La situation de discours pur s'inscrit en effet dans l'*espace-temps*, et le non-ici peut renvoyer à un non-maintenant comme l'attestent les exemples suivants avec *sortir*. En tant qu'achèvement (cf. note 6) *sortir* donne lieu à deux (voire trois) interprétations temporelles opposées tributaires de la préposition : « Léa sort *en* boîte. » (procès futur ou habituel) et « Léa sort *de* boîte. » (procès passé). Dans chacune de ces interprétations, le *temps de dicto* permet l'utilisation du Présent. Il en est de même dans un énoncé dépourvu d'indication spatiotemporelle explicite comme « Il répare *les* voitures. », où selon toute vraisemblance, la marque de pluriel indique une répartition *dans le temps* qui diverge de l'uni(cité) temporelle de « Il répare *la* voiture. ». Dans ces exemples, si le Présent ne permet pas d'asserter que l'évènement dénoté par le sémantisme de la lexis est contemporain du moment d'énonciation (*temps de re*), il actualise en revanche le *temps de dicto*.

En dépit de sa position d'observateur inhérente au travail de description (perspective de décodage), comme le montre l'influence de l'implication 1 dans le comportement temporel du Présent, le linguiste doit également envisager les stratégies d'encodage de l'information.

¹² Il est probable que le Présent dit « narratif » incarne lui aussi une forme d'actualisation. Cette question demandant à être approfondie, nous l'avons écartée faute de place.

Enfin, cette analyse montre que, même s'il paraît souvent commode de qualifier le Présent de « gnomique », « historique », « panchronique », etc., ces qualités ne découlent nullement du sémantisme du tiroir verbal mais de celui de sa lexis ou du contexte pragmatique dans lequel il apparaît.

Références

- ADAM, Jean-Michel (1992). *Les textes : types et prototypes. Récit, description, argumentation, explication et dialogue*. Paris : Nathan.
- ADAM, Jean-Michel, LUGRIN, Gilles & REVAZ, Françoise (1998). Pour en finir avec le couple récit/discours, *Pratiques 100 - Les Temps verbaux*, 81-98.
- BENVENISTE, Émile (1974). *Problèmes de linguistique générale 2*. Paris : Gallimard.
- CHEVALIER, Jean-Claude (1996). De Guillaume à une linguistique du signifiant, *Modèles linguistiques 33 - Temps et langage (III)/L'invariant*, 77-92.
- COMRIE, Bernard (1985). *Tense*. Cambridge : Cambridge University Press.
- CULIOLI, Antoine (1990). *Pour une linguistique de l'énonciation. Opérations et représentations, Tome 1*. Paris/Gap : Ophrys.
- CULIOLI, Antoine (1999). *Pour une linguistique de l'énonciation. Formalisation et opérations de repérage, Tome 2*. Paris/Gap : Ophrys.
- FAUCONNIER, Gilles (1984). *Espaces mentaux*. Paris : Minuit.
- FRAISSE, Paul (1957/1967). *Psychologie du temps*, Paris : Presses Universitaires de France.
- FUCHS, Catherine (1991). Les typologies de procès : un carrefour théorique interdisciplinaire, in *Les Typologies de procès*, FUCHS Catherine (Ed). Paris : Klincksieck, 9-17.
- GOSSELIN, Laurent (1996). *Sémantique de la temporalité en français*. Louvain-la-Neuve : Duculot.
- KAHN, Félix (1954). *Le système des temps de l'indicatif*. Genève : Droz.
- KLEIN, Wolfgang (1994). *Time in language*. London: Routledge.
- LE GOFFIC, Pierre & LAB, Frédérique (2000). Le présent « pro futuro », *Cahiers Chronos 7 - Le présent en français*, 77-98.
- MARTIN, Robert (1985). Langage et temps de dicto, *Langue française 67 - La pragmatique des temps verbaux*, 23-37.
- ORIGGI, Gloria & SPERBER, Dan (2000). Evolution, communication and the proper function of language, in *Evolution and the Human Mind: Language, Modularity and Social Cognition*, CARRUTHERS Peter & CHAMBERLAIN Andrew (Eds). Cambridge : Cambridge University Press, 140-169.
- PÖPPEL, Ernst (1997). A hierarchical model of temporal perception, *Trends in Cognitive Sciences 1- 2*, 56-61.
- PÖPPEL, Ernst (2004). Lost in time: a historical frame, elementary processing units and the 3-second window, *Acta Neurobiologiae Experimentalis 64-3*, 295-301.
- REICHENBACH, Hans (1947). *Elements of Symbolic Logic*. New York/London : The Free Press/Collier-Macmillan.
- SERBAT, Guy (1988). Le prétendu « présent » de l'indicatif: une forme non déictique du verbe. *L'information grammaticale 38 - Temps verbaux et temporalité*, 32-35.
- TOURATIER, Christian (1996). *Le système verbal français*. Paris : Armand Colin.
- VENDLER, Zeno (1957). Verbs and Times, *Philosophical Review 66*, 143-160.
- VETTERS, Carl (1996). *Temps, aspect et narration*. Amsterdam : Rodopi.

LA VARIATION DIATOPIQUE DU FRANÇAIS DANS L'ŒUVRE LITTÉRAIRE D'YVES VIOLLIER (VENDÉE)

Inka WISSNER

inkawissner@gmx.net

Université de Bonn

Am Hof 1- 53113 Bonn

Allemagne

et Université Paris-Sorbonne

1 rue Victor Cousin - 75230 Paris cedex 05

Abstract : In the field of contemporary French regional variation, this article presents an ongoing research project on the basis of a literary corpus, the rich works of the novelist Yves Viollier (Vendée, France). Definitions and methodological issues are addressed, followed by the analysis of three sample variables (*dail*, *noceur*, *attendre à*) to illustrate their nature as well as their textual particularities in the literary context.

1. Le français dans sa variation diatopique

Comme l'annonce le titre, l'étude en cours présentée dans cet article se situe dans le domaine de l'analyse de la variation diatopique du français. Après un bref aperçu du concept de la variation du français au niveau de la diatopie, suivra une description de la méthodologie mise en œuvre pour aborder les textes littéraires de l'écrivain contemporain Yves Viollier qui ont été choisis comme corpus. Parmi les nombreux particularismes linguistiques identifiés, je choisirai trois éléments à titre d'exemples : les termes *dail* et *noceur*, et la construction *attendre quelqu'un/ quelque chose à* + infinitif.

Selon l'approche variationniste, que nous défendons ici, toute langue est marquée par la variation : une langue standard, qui s'est implantée au cours de l'histoire par son expansion et sa codification, comporte bien sûr aussi une diversification (sociale, régionale et contextuelle). La *variation diatopique*, c'est-à-dire dans l'espace, lui est donc inhérente¹.

Pour désigner l'ensemble des usages du français qui ont cours dans une aire linguistique donnée, ou qui linguistiquement distinguent une aire donnée de la variété considérée comme normative, nous parlons ici de différentes *variétés diatopiques* du français (qui sont souvent appelées aussi *français régionaux* par commodité d'expression). Les *diatopismes*, "fait[s] de langue géographiquement marqué[s]" (Rézeau, 2001 : 16, ici DRF) sont généralement décrits par rapport au français dit de référence, tel qu'il se présente sans marque(s) dans les ouvrages normatifs (dictionnaires et grammaires usuels). Un diatopisme est loin d'être nécessairement restreint à une région, comme pourrait l'évoquer le terme *régionalisme*². De plus, il peut

¹ Déjà E. Coseriu, en décrivant la différenciation interne des langues romanes (illustrée entre autres par l'exemple du français) souligne que dans sa structure, une langue historique varie sur le plan *diastatique* (socioculturel), *diatopique* (dans l'espace géographique, v. *topos* gr. "lieu") et *diaphasique* (selon les circonstances, en fonction de modalités d'expression) (Coseriu, 1981 : 12, 15). Pour la notion de variation diatopique dans le standard, v. surtout J.-P. Chambon (2005 : 6-9). Le discours contemporain autour de la standardisation des langues doit beaucoup à E. Haugen, qui souligne que l'évolution d'une langue standard (ou en voie de standardisation) implique les processus de « selection, codification, acceptance, and elaboration » (1966 : 933, v. aussi 922).

² Couramment employé depuis quelques décennies comme terme générique, le terme *régionalisme* évoque en effet la notion d'une limitation d'emploi dans l'espace d'un phénomène linguistique donné

apparaître dans le discours parlé et écrit, de tous les milieux, et être d'extension géographique fortement variable. L'inconscience de son usage dans la représentation des usagers (aspect souvent discuté et avancé comme trait définitoire du diatopisme), comme la conscience de son usage, dépendent de la nature du phénomène et du locuteur (v. p.ex. Tuaillon, 1983 : 23-24 et Rézeau, 1984 : 20-21), ainsi que sa stigmatisation par la société.

Au niveau de la nature des faits diatopiques, nous pouvons distinguer non seulement les diatopismes lexicaux, sémantiques (p.ex. Rézeau, 1984 : 11-13) et morphosyntaxiques (déjà Schmitt, 1977), mais aussi lexématiques, phraséologiques, pragmatiques, phonétiques et phonologiques, où s'ajoutent aussi les diatopismes de statut (Thibault, 1996 : 334-365). En effet, un fait diatopique peut être rare dans une communauté linguistique et tout à fait usuel dans une autre. Ceci est le cas de l'emploi de l'article défini devant les prénoms (*le Laurent, la Chantal*), par exemple, qui est nettement plus fréquent dans certaines régions que dans d'autres, sans la connotation péjorative qui lui est attribuée en français de référence³. La fréquence plus ou moins élevée d'un diatopisme étant liée à son fonctionnement syntagmatique et paradigmatique, les particularismes peuvent concourir librement avec un équivalent général ou bien être les seuls à la disposition des locuteurs pour exprimer une notion donnée (v. p.ex. Rézeau, 1984 : 15-16, 19 ou Tuaillon, 1977 : 18-19).

En même temps, les diatopismes sont aussi généralement catégorisés selon leur origine, et ceci depuis le début du 20^e siècle (v. p. ex. Chambon, 2005 : 14). Nous pouvons avoir affaire à des **archaïsmes**, i.e. à des survivances du français (depuis l'ancien français), qui demeurent généralement dans des régions éparses et plutôt éloignées de Paris et des grands centres urbains. Ainsi, la fameuse triade *déjeuner/dîner/souper* coexiste (encore) avec le français général *petit-déjeuner/déjeuner/dîner* notamment pour de nombreux francophones d'Europe et du Québec (v. aussi l'interjection *Dame !*, maintenue depuis le 17^e siècle dans certaines régions de France, DRF : 339-340). Les **innovations (internes) dues au dynamisme propre de la langue**, au contraire, se diffusent le plus souvent à partir de villes jouissant d'une position socioculturelle/politique supérieure dans une zone d'influence donnée (en général vers des centres urbains d'abord, puis – aussi à partir de là – vers les lieux plus éloignés). Ainsi, si nous prenons l'exemple de la désignation *bugne* pour la pâtisserie traditionnelle du Mardi Gras, elle est attestée à Lyon depuis le 16^e siècle et est aujourd'hui courante dans une très large aire de l'Est de la France (DRF : 170-172). Pour décrire ce phénomène de diffusion, les chercheurs étudiant les variétés diatopiques de l'anglais (en *dialectology*, à ne pas confondre avec la dialectologie française), parlent aussi de *contagious and hierarchical dialect diffusion* (Wolfram *et al.*, 2006 : 153-155). Les **innovations par emprunt**, enfin, sont ces particularismes intégrés dans une variété donnée par transfert linguistique à partir

(p.ex. Tuaillon, 1983 : 2, Rézeau, 1984 : 9-10). Par contre, il est plutôt adapté pour désigner les particularismes *propres* à une région. Lorsqu'on observe le français à l'échelle francophone, ce terme est moins approprié pour désigner les faits qui sont également attestés en dehors d'une aire linguistique continue, puisqu'ils peuvent dessiner des configurations aérologiques diverses.

³ Le phénomène de l'emploi de l'article devant les noms propres a déjà provoqué de nombreuses discussions et remarques. Si certains le classifient de 'rural', en référence à un village isérois (Tuaillon, 1983 : 4), d'autres soulignent que le marquage 'populaire' serait plus adapté (Rézeau, 1986b : 45). Seule une enquête systématique couvrant la totalité de l'hexagone, et si possible de la francophonie, permettrait de mesurer plus précisément l'emploi de ce phénomène. Une telle étude faisant défaut, des enquêtes ponctuelles menées dans la première moitié de 2006 auprès d'une vingtaine de locuteurs francophones d'origines différentes m'ont permis de constater des différences d'emploi non négligeables en Europe : dans des milieux et des contextes comparables, le phénomène est aujourd'hui plutôt rare p.ex. à Lyon, en Auvergne, en Vendée, à Neuchâtel ou dans le canton de Vaud (Suisse), mais stylistiquement neutre dans le Jura suisse (Thibault, 1997 : 802) ou d'une fréquence remarquable dans différents milieux à Nancy.

d'autres langues et variétés linguistiques (ainsi l'Allemand ou l'Alsacien, ou des variétés gallo-romanes comme les dialectes en France). Après un processus d'adaptation (lexicale, sémantique, phonétique et/ou morphosyntaxique) plus ou moins marqué, en fonction des habitudes langagières des locuteurs, de leur proximité linguistique à la variété emprunteuse et du contexte dans lequel a lieu le transfert, le transfert peut se faire directement d'une variété à une autre ou passer par l'intermédiaire d'(une) autre(s) variété(s), d'un registre linguistique ou d'un milieu socioculturel particuliers (v. aussi Thibault, 1996 : 364). Ainsi, dans l'Ouest de la France, le français emprunte des diatopismes p.ex. au languedocien, par l'intermédiaire du dialecte, comme *mogette* (DRF : 686-688 ; v. ci-dessous).

La catégorisation canonique des diatopismes selon leur origine est en effet à préciser par une description plus détaillée, tenant également compte de la diffusion des diatopismes et de leur dimension sociale (v. p.ex. Chambon, 2005 : 12-13). On peut observer différentes sous-normes (locales, régionales) à côté de la/des norme(s) nationale(s), qui varient en fonction des phénomènes, des aires linguistiques et des milieux (*op.cité* : 10, 14-15). Si la pression sociale exercée depuis les couches 'supérieures' (normes internationale, nationale ou régionale) est le pôle dominant, il y a aussi un prestige 'caché', les locuteurs ayant tendance à s'adapter au langage de leur communauté linguistique comme l'indique le terme *covert prestige* du sociolinguiste W. Labov (Labov, 1994 : 23-24 et Chambers *et al.*, 1998 : 85). Ce besoin d'être perçu favorablement par le groupe identitaire (social et/ou local/régional) encourage évidemment aussi l'emploi de diatopismes, dont certains peuvent être perçus comme familiers ou populaires, d'autres comme tout à fait normatifs.

2. La variation diatopique et la littérature contemporaine

2.1. *Le corpus : l'œuvre du Vendéen Yves Viollier*

Pour analyser le français contemporain dans sa variation diatopique, j'ai choisi l'œuvre littéraire contemporaine d'un écrivain de Vendée (Centre-Ouest de la France) : l'ancien professeur de français Y. Viollier est né près de La Roche-sur-Yon en 1946 et continue d'y écrire. Son œuvre, particulièrement complexe et riche avec 24 ouvrages à l'heure actuelle (publiés pour la plupart chez R. Laffont), s'étale sur quatre décennies. Les observations et exemples présentés dans cet article résultent du dépouillement d'ouvrages de quatre phases littéraires différentes de l'auteur étudié (Viollier, 1975 ; 1982 ; 1994 ; 2002)⁴. Dans ses ouvrages presque toujours ancrés dans sa région natale et écrits pour un large public francophone, le Vendéen qui se considère comme un écrivain populaire 'du réel' (cf. notre entretien avec l'auteur de 2006), invite les lecteurs indigènes à se sentir confirmés dans leur identité régionale, et donne surtout aux lecteurs exogènes la possibilité de connaître, à travers son regard artistique, sa vision du monde – dont fait aussi partie sa langue.

2.2. *L'approche différentielle : identifier les diatopismes*

Pour analyser le français contemporain dans sa variation diatopique, mon projet s'appuie sur des méthodes et des matériaux fiables établis dans les travaux de lexicographie différentielle surtout par les contributeurs des grands dictionnaires diatopiques (Thibault, 1997 ; Poirier, 1998 ; DRF). L'identification des diatopismes dans l'œuvre d'Y. Viollier se fait en partant du cadre du français de référence, i.e. du corpus d'exclusion (v. Gleßgen *et al.*, 2005 : XIV et déjà Robez-Ferraris, 1988 : 16) composé des dictionnaires courants – surtout du *Trésor de la*

⁴ Après un début de carrière poétique dans les années 1970 (Viollier, 1975), le romancier adhère au mouvement dit 'régionaliste' surtout dans les années 1980 (Viollier, 1982) et prend aussitôt du recul – ceci aussi sous une direction littéraire à tendance plutôt jacobine (Viollier, 1994) ; au 21^e siècle, il revendique son identité régionale avec plus de rigueur sans pourtant puiser dans le stock du dialecte, en tâchant de rester fidèle au français tel qu'il est d'usage en Vendée (Viollier, 2002).

langue française informatisé (Atilf, 2004, ici TLFi) – et du *Bon Usage* pour la grammaire (Grevisse, 1993). Le *corpus d'inclusion*, comme je propose de le dénommer, est constitué du DRF et de ses matériaux préparatoires manuscrits, du *Französisches etymologisches Wörterbuch* de W. von Wartburg (ici FEW), ainsi que de son *Fonds documentaire*, ouvrages décrivant les variétés linguistiques du Centre-Ouest de la France notamment au niveau dialectal⁵. Ces recherches à base d'outils textuels sont complétées d'enquêtes orales que nous avons menées en juillet 2006 avec des locuteurs vendéens (échantillon de 13 informateurs).

2.3. *Les diatopismes dans leur contexte littéraire*

La littérature, source lexicographique traditionnelle, est en effet dépouillée avec profit depuis que les chercheurs étudient la variation diatopique du français (v. Rézeau, 1986a et DRF), et certains se sont déjà intéressés à la place particulière des diatopismes dans l'œuvre littéraire (p.ex. Robez-Ferraris, 1988 ou Franconie, 1991). Néanmoins, à l'heure actuelle, une analyse de l'ancrage pragmatique des diatopismes dans la littérature hexagonale fait encore défaut ; c'est ainsi que la présente étude tâche de combler un vide en décrivant les éléments diatopiques dans l'œuvre d'Y. Viollier dans leur contexte littéraire, avec les particularités textuelles qui les entourent. En partant des dépouillements entrepris, je propose de regrouper les divers particularismes discursifs qui accompagnent les diatopismes en trois catégories : les attributions aux niveaux textuels divers à l'intérieur du discours, le métalangage de l'auteur, et les fonctions textuelles que remplissent les éléments diatopiques.

En observant les occurrences des particularismes diatopiques en fonction des différents niveaux textuels, on constate que les diatopismes peuvent être attribués (plus ou moins consciemment) au discours du narrateur, au discours rapporté, au discours de certains personnages particuliers, ou p.ex. au récit sur le passé ou le présent⁶. Ces attributions s'expliquent souvent simplement par des besoins d'expression ponctuels dans un contexte précis ; en même temps, elles sont aussi le résultat de stratégies discursives de la part de l'écrivain, en fonction de l'image qu'il désire donner de son écriture. Ainsi, si Y. Viollier n'hésite pas à inclure des diatopismes dans toutes les parties du discours dans les années 1980 (id. pour des éléments du dialecte, évités par la suite), il tâche de restreindre tout diatopisme au discours rapporté dans les années 1990 (sans vouloir stigmatiser les personnages), entre autres afin de répondre aux desiderata d'un directeur littéraire favorisant la langue jacobine. Depuis le 21^e siècle, compte tenu d'une réception (plus) positive des diatopismes dans son œuvre par son nouveau directeur littéraire, les lecteurs et la critique, l'auteur recourt aux diatopismes dans le discours rapporté *et* dans la narration, quand il s'agit, pour lui, du mot « juste » et « irréversible » (cf. entretien avec l'auteur).

Dans la littérature diatopique adressée à un public francophone contemporain, les auteurs intègrent également des éléments métalinguistiques qui accompagnent des particularismes linguistiques comme chez J. Anglade, H. Vincenot ou J.-P. Chabrol (v. DRF). Ce discours peut se présenter sous forme de gloses en bas de page (Viollier, 1982), intégrées dans le texte

⁵ La bipartition *corpus d'exclusion* vs *inclusion* a pour but de séparer deux types de sources selon leurs fonctions primaires dans le cadre de mon analyse. Les dictionnaires et grammaires normatifs peuvent aussi être des sources valables pour la variation diatopique, et les ouvrages catégorisés dans le corpus d'inclusion ne se limitent bien sûr pas à une description des diatopismes du français.

⁶ Dans le discours romanesque, M. Bakhtine souligne l'importance de distinguer les unités compositionnelles du discours, i.e. le discours de l'auteur, des différents types de narrateur, des personnages, et les formes littéraires et genres non-littéraires intégrées dans le roman (1978 : 87-89). Selon le philosophe russe et théoricien du roman, l'échange des 'langues' (parlers des groupes 'sociaux' et 'territoriaux', associés à des 'points de vue' dans un sens idéologique), est une des caractéristiques propres au roman contemporain (op. cité : 134).

(soit discrètement dans le fil du récit, stratégie à laquelle recourt généralement Y. Viollier, soit entre parenthèses), comme glossaires en fin d'ouvrage ou comme gloses dans un appareil critique de l'éditeur (guère chez notre auteur). Ces gloses pouvant se présenter sous forme par exemple d'explications, de synonymes ou de métonymes, elles ne sont généralement pas présentées dans une perspective ou avec une exigence lexicographiques : elles sont incluses pour faciliter le décodage du texte, et afin 'd'apporter un plus' de création littéraire (entretien avec l'auteur). Par l'emploi de marqueurs typographiques modifiant les diatopismes (emploi de guillemets, représentation graphique en italique/gras), l'auteur peut recourir à un tour, tout en l'attribuant à une communauté linguistique ou à un personnage décrits dans le récit, en s'en distanciant plus ou moins explicitement. En outre, des commentaires du genre « comme on le dit chez nous » accompagnent certains diatopismes, soulignant ainsi leur fonction identitaire. Ces stratégies discursives sont l'indice de la conscience qu'a l'auteur de la régionalité d'un fait de langue, ou du moins de probables difficultés de compréhension.

Dans une vue fonctionnelle, il importe également de souligner les fonctions textuelles diverses qu'ont les diatopismes dans le discours littéraire. Comme le remarque H. Franconie, la signification des diatopismes varie selon le contexte diamésique, et leurs « effets expressifs sont éphémères dans le discours parlé, liés au seul moment de la communication, alors qu'ils sont rémanents dans une œuvre écrite » (Franconie, 1991 : 48). Surtout dans les romans avec une forte présence de la province, l'emploi de diatopismes sert à exprimer une idée ou une notion précise, à évoquer des stéréotypes, à mettre en valeur un terme ou une situation, à décrire des *realia* propres à la culture ou la vie régionales, dans des domaines comme l'environnement, l'habitat traditionnel, les outils ou la cuisine (v. p.ex. Rézeau, 1984 : 10-11 ou Tuailon, 1977 : 19) ou simplement à augmenter l'authenticité du discours.

2.4. *L'œuvre d'Yves Viollier : un choix d'exemples*

Afin d'illustrer la variation diatopique dans les écrits d'Y. Viollier, nous pourrions traiter de nombreux cas de diatopismes, surtout au niveau lexical. La plus grande partie relève des champs sémantiques du travail (notamment agricole), de la géographie décrivant le pays, de la flore, de la famille et de la nourriture, mais également simplement du vocabulaire courant.

Ainsi nous trouvons des diatopismes comme *carreau* (pour *carré*) (de chocolat) ou *mojette* s.f. "haricot vendéen blanc et sec (parfois vert)", du champ sémantique de la nourriture ; *vergne* s.m. pour le fr. de réf. *aulne* s.m., du domaine de la flore ; *dail* s.m. pour *faux* s.f., ou *toucher* v. tr. "conduire un attelage de bœufs ou de vaches", de l'agriculture ; *nénène* s.f. (fam.) comme terme affectueux pour *marraine* s.f., ou *noceur* s.m. au sens "invité à un mariage", relevant de la famille et de la vie sociale. Mais d'autres relèvent du lexique quotidien de domaines divers, comme *attendre quelqu'un/quelque chose* à suivi de l'infinitif, pour *attendre que quelqu'un/quelque chose* suivi du subjonctif, *s'emporter* (la peau) v. pron. pour le fr. de réf. *s'écorcher* (la peau), ou *hier (au) tantôt* pour le fr. de réf. *hier après-midi*.

Les sujets traités par l'auteur dans ses ouvrages, ses préférences thématiques (la nature et la vie quotidienne des vendéens) et l'usage stratégique de diatopismes évoquant 'l'authentique' sont en effet à l'origine d'une partie de cette distribution lexicale. Mais les domaines liés au quotidien, à l'affectif et aux pratiques techniques éventuellement propres à une région comportent naturellement un taux de diatopismes relativement élevé en général, à l'écrit et à l'oral (v. p.ex. DRF : 8). Dans notre corpus, de tels diatopismes lexicaux sont largement plus fréquents que les diatopismes morphosyntaxiques et phonétiques, qui ne dépassent pas la dizaine (pour une proportion similaire à titre de comparaison, v. l'étude sur le français parlé à Vourey en Isère : Tuailon, 1983 : 370). Tandis que les diatopismes phonétiques sont rares à l'écrit (chez Y. Viollier l'orthographe *not'maître* évoquant la chute du [R] final dans le discours rapporté fait exception : Viollier, 1994 : 314), le nombre d'occurrences par particularisme morphosyntaxique peut être très élevé dans l'œuvre étudiée, divergeant

vraisemblablement peu du discours parlé (si l'on accepte par exemple le datif éthique comme diatopisme de statut).

Après ce bref aperçu d'ensemble, nous nous concentrerons sur trois exemples de diatopismes, en indiquant quelques aspects choisis qui caractérisent notre analyse en cours. Le premier diatopisme qui nous intéresse relève du domaine grammatical et serait à présenter, dans une monographie, dans une section consacrée aux aspects morphosyntaxiques, séparée d'un glossaire. Pour évoquer qu'un groupe d'amis se rassemble un matin une fois que l'humidité dans l'air s'est dissipée, le narrateur écrit :

« On avait attendu la rosée à fondre⁷ pour se rejoindre sur le chantier. »
(Viollier, 1982 : 75).

Le verbe transitif de cet énoncé, à lemmatiser sous sa forme canonique *attendre quelqu'un/quelque chose* à + infinitif, est un exemple d'un diatopisme morphosyntaxique, en concurrence avec le français de référence *attendre que quelqu'un/quelque chose* + subjonctif. Absent des dictionnaires généraux et différentiels, sa vitalité est pourtant bien attestée dans les enquêtes préparatoires du DRF (Rézeau, 1994-96) de même que dans les enquêtes menées par moi-même (d'usage dit 'courant'). Le deuxième particularisme étudié, lui, ne diverge pas du français de référence au niveau de la forme, mais du sens :

« Germain passa prendre les noceurs au volant de sa camionnette. » « Les bouteilles circulèrent dans les rangs des noceurs. » (Viollier, 1994 : 107, 110).

Avec le terme *noceur*, nous n'avons pas ici affaire au sens du français général "fêtarde, personne qui aime faire la fête", même si celui-ci est également courant en Vendée. Il s'agit d'un diatopisme sémantique *noceur* "invité à un mariage" (généralement employé au pluriel), bien attesté en Vendée aujourd'hui ('courant' d'après mes enquêtes). Si l'on observe des variantes dialectales du type lexical, on constate que quatre glossaires régionaux en enregistrent plusieurs pour la Vendée⁸ ; et selon von Wartburg, des variantes diverses (ainsi *noceux*, *-euse*, *noçoux*, *-ouse*) sont attestées dans le grand Centre-Ouest de la France, une variante au Canada (*noceux*), une en Suisse et une dans les Pyrénées (FEW VII : 244a s.v. *NUPTIAE*) (les deux dernières attestations sont isolées). Avec l'implantation dans les dialectes d'une aire compacte dans le grand Ouest de la France, il ne s'agit pas d'une survivance du français commun qui se serait conservée dans des régions éparses, mais d'une innovation en français. Vu que nous avons une attestation en Amérique du Nord, où le français a été principalement importé au 18^e siècle à partir du grand Ouest de la France, nous pouvons présumer que le type lexical a déjà été suffisamment fréquent en français dans le grand Ouest hexagonal à l'époque classique. En tenant également compte de la formation de mot qui est bien française : v. la suffixation en *-eur* (et non pas glissement de sens à partir de l'homonyme du français général), il s'agit le plus probablement d'une innovation due au dynamisme propre du français. La variante *noceur* s'est probablement développée dans un milieu français dans le Centre-Ouest pour ensuite avoir pénétré les dialectes. Théoriquement, cette formation peut en effet avoir lieu n'importe où dans la francophonie (v. les attestations dites 'isolées').

⁷ L'emploi du v. *fondre* pour fr. gén. *se dissiper* en parlant de la rosée est de caractère idiosyncratique : il n'est confirmé ni par les ouvrages de référence (p.ex. TLFi, Rey, 2001) ou différentiels consultés (p.ex. Rézeau, 1984, 1994-96, 1999-2000, DRF), ni par mes enquêtes en Vendée.

⁸ Marais vendéen *nosur* s.m., *-uz* s.f. (Svenson, 1959 : s.v. *nosur*) ; Vendée *noçour*, *-ouse* s.m. et f., *noçur* (BéaParlange, 1983 : 35) ; Vendée *noçours* "noceurs, gens de la noce" (Gautier, 1996 : 180-181) ; Triaize *noçour* s.m. "noceur, invité à une noce", *-ouse* s.f. (Pivetea, 1987 : s.v. *noçour*).

Observons en dernier lieu un diatopisme lexématique le plus probablement emprunté à l'occitan :

« [...] ils avaient acheté ensemble une faucheuse pour les foins, avec l'autorisation de leurs mères. C'était la première qui rentrait dans la commune. Elle fit hausser les épaules à plus d'un vieux porteur de dail [i.e. faux] : une mécanique ! [...] il fallait que ces jeunes soient complètement fous ou qu'ils soient vraiment devenus fainéants pour ne pas avoir le courage de couper à la main comme tout le monde ! » (Viollier, 1982 : 71)

Le diatopisme *dail* "instrument tranchant, formé d'une lame arquée fixée à un long manche, que l'on manie à deux mains pour couper l'herbe (autrefois du blé)" (issu d'une forme ancienne gallo-romane du sud, de lat. DACULUM : v. FEW III, 2b-3a s.v. *DACULUS, -A) est décrit en détail dans le DRF, du fait de son extension large dans une grande partie du sud de la France, allant de la Loire à l'Ouest jusqu'au Jura à l'Est (DRF : 337-338). Je voudrais attirer ici l'attention sur quelques particularités pragmatiques et textuelles, et les mettre en relation avec les autres emplois trouvés dans notre corpus jusqu'à présent. Dans le passage cité, où le diatopisme *dail* apparaît la 1^{ère} fois dans ce roman régionaliste, il est accompagné d'une glose en bas de page : elle est constituée de l'indication du synonyme de référence, *faux*, avec lequel le diatopisme est en concurrence libre, au niveau textuel externe (en français, en l'occurrence en Vendée) et interne (dans l'œuvre étudiée). Comme l'indique la glose, l'auteur est conscient de son caractère éventuellement incompréhensible pour le lecteur exogène. Y. Viollier recourt à ce diatopisme d'une part pour des raisons stylistiques afin d'éviter des répétitions, et ceci à deux reprises (Viollier, 1982 : 74/75 ; Viollier, 1994 : 245), d'autre part pour mieux évoquer la réalité rurale telle qu'elle se présente à l'époque décrite.

Pour ce qui est des attributions discursives, on constate que le terme peut apparaître dans le discours d'un paysan (Viollier, 1994 : 245). Pourtant, le terme *dail* n'est pas attribué à la langue parlée, vu son apparition fréquente dans la narration même (p.ex. Viollier, 1982 : 71, 74-77), contre quelques emplois dans le discours direct (p.ex. Viollier, 1982 : 72/73 ; 1994 : 245). Par contre, il est à retenir que tous les emplois rassemblés à l'heure actuelle apparaissent dans des récits situés au début du 20^e siècle. Peut-être ceci n'est-il lié qu'au fait que le contexte se prête à décrire cet objet agricole ; néanmoins, une éventuelle attribution consciente au discours du passé reste encore à vérifier lors des dépouillements à suivre.

3. Conclusion

Après ces quelques aspects choisis concernant le français dans sa variation diatopique et une analyse exemplaire d'un corpus littéraire d'aujourd'hui, je voudrais, en guise de conclusion, en venir à l'intérêt d'une étude du type de celle que nous venons d'entrevoir. Si nous prenons le point de vue linguistique, ce travail apporte d'autres éléments au réseau de données à la disposition des scientifiques décrivant le français contemporain. L'ancrage pragmatique et l'étude des particularités textuelles pourraient également intéresser des chercheurs d'autres domaines, ainsi de la pragmatique textuelle ou de la littérature. Pour le monde de l'édition et le lecteur, une analyse approfondie des diatopismes dans une œuvre littéraire peut servir de base pour une vulgarisation sous forme de glossaire (v. Thibault, 2006 : 170). A long terme, ce genre d'étude pourra contribuer aussi à éclairer l'inhérence de la variation (sociale, contextuelle et diatopique) à la langue française, malgré les idéologies du standard qui peuvent y voir l'échec de tout effort d'unification linguistique.

Références bibliographiques

- ATILF (2004). *TLFi : Trésor de la langue française informatisé*. Nancy : CNRS Éditions.
 BAKHTINE, Michail (1978). *Esthétique et théorie du roman*. Paris : Gallimard.
 BÉAPARLANGE (1983). *Vie mentale, affective, sociale*. La Roche-sur-Yon : Arantèle.

- CHAMBERS, Jack & TRUDGILL, Peter (1998). *Dialectology*. Cambridge : Cambridge UP.
- CHAMBON, Jean-Pierre (2005). Après le *Dictionnaire des régionalismes de France* : bilan et perspectives, in *La lexicographie différentielle du français et le DRF*, Gleßgen, Martin-D. & Thibault, André (Eds). Strasbourg : PU Strasbourg, 3-29.
- COSERIU, Eugenio (1981). Los conceptos de "dialecto", "nivel" y "estilo de lengua" y el sentido propio de la dialectología, *Lingüística española actual*, 3 : 1-32.
- FRANCONIE, Hélène (1991). Spécificité du régionalisme écrit, in *Parlers et cultures : le français régional*, Pelen, Jean-Noël & Martel, Claude (Eds). Provence : PU Provence, 46-51.
- GAUTIER, Michel (1996). *Vu dire. Chroniques vendéennes*, Mougou : Geste Éditions.
- GLEßGEN, Martin-D. & THIBAUT, André (2005). La 'régionalité linguistique' dans la Romania et en français, in id. (Eds). Strasbourg : PU Strasbourg, III-XVII.
- GREVISSE, Maurice (1993). *Le Bon Usage*. Refondu par André Goosse. Paris : Duculot.
- HAUGEN, Einar (1966). Dialect, Language, Nation, *American Anthropologist*, 68 : 922-935.
- LABOV, William (1994). *Principles of Linguistic Change*, vol. I. Oxford : Blackwell.
- PIVETEA, Vianney (1987). *Parlange Marouin. Lexique poitevin-français*. Fiche. s.ed.
- POIRIER, Claude (1998). *Dictionnaire historique du français québécois*. Sainte-Foy : PU Laval.
- REY, Alain (2001). *Le Grand Robert de la langue française*. Paris : Le Robert/Sejer.
- RÉZEAU, Pierre (1984). *Dictionnaire des régionalismes de l'Ouest entre Loire et Gironde*, Les Sables-d'Olonne : Le Cercle d'Or.
- RÉZEAU, Pierre (1986a). *Bibliographie des régionalismes du français*. Paris : Klincksieck.
- RÉZEAU, Pierre (1986b). Les régionalismes et les dictionnaires du français, in *La lexicographie québécoise*, Boisvert, Lionel & Poirier, Claude (Eds). Québec : PU Laval, 41-52.
- RÉZEAU, Pierre (1994-1996). EnqDRF 1994-96 ms. : Manuscrits d'enquêtes. Nancy : Atilf.
- RÉZEAU, Pierre (1999-2000). DRF : Manuscrits en fichier informatique. Nancy : Atilf.
- RÉZEAU, Pierre (2001). *Dictionnaire des régionalismes de France (DRF). Géographie et histoire d'un patrimoine linguistique*. Bruxelles : Duculot.
- ROBEZ-FERRARIS, Jacqueline (1988). *Les richesses du lexique d'Henri Vincenot, auteur Bourguignon*. Paris : Klincksieck.
- SCHMITT, Christian (1977). La grammaire des 16^e et 17^e siècles et les langues régionales, *Travaux de Linguistique et de Littérature*, 15 : 215-225.
- SVENSON, Lars-Owe (1959). *Les parlers du Marais Vendéen*. Göteborg : E. B. Aktiebolag.
- THIBAUT, André (1996). Québécoismes et helvétismes : éclairages réciproques, in *Français du Canada – français de France*, Lavoie, Thomas (Ed). Tübingen : Niemeyer, 333-376.
- THIBAUT, André (1997). *Dictionnaire suisse romand*. Genève : Zoé.
- THIBAUT, André (2006). Glossaristique et littérature francophone, *Revue de Linguistique Romane*, 70 : 143-179.
- TUAILLON, Gaston (1977). Réflexions sur le français régional, *Travaux de Linguistique et de Littérature*, 15 : 7-29.
- TUAILLON, Gaston (1983). *Les régionalismes du français parlé à Vourey, village dauphinois*. Paris : Klincksieck.
- VIOLLIER, Yves (1975). *Les noces de Claudine*. Les Sables-d'Olonne : Le Cercle d'Or.
- VIOLLIER, Yves (1982). *La cabane à Satan*. Paris : Robert Laffont.
- VIOLLIER, Yves (1994). *Les pêches de vigne*. Paris : Robert Laffont.
- VIOLLIER, Yves (2002). *Les sœurs Robin*. Paris : Robert Laffont.
- WARTBURG, Walther von (1922-2004). *Französisches etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes*, vol. I-XXV. Bâle et al. : Zbinden et al.
- WOLFRAM, Walt & SCHILLING-ESTES, Natalie (2006). *American English*. Malden : Blackwell.

PSYCHOLINGUISTIQUE ET DÉVELOPPEMENT LANGAGIER

TERMES DE COULEUR EN DISCOURS ET CONCEPTS DE COULEUR

Caroline CANCE

ccance@free.fr

LCPE-LAM/IJLRD, 11 rue de Lourmel

75015 Paris

Abstract : Our work focuses on the relationship between colour perception, cognition and representations in discourse. Linguistic analyses were carried out in order to study the co-construction of cognitive representations, cultural knowledge and lexicalisation of colours. Three colour terms used in various discourses were collected in relation to different types of sensory experiences. Their analysis shows that colour terms may relate to various dynamic cognitive constructions such as individual and shared knowledge.

1. Introduction

Dans un contexte scientifique faisant la part belle aux Sciences de la Cognition, l'intérêt pour la question de l'articulation entre langage, cognition et perception est renouvelé. La confrontation de ce questionnement à une demande industrielle¹ - *comment se construit le jugement perceptif des usagers à propos d'habitacles automobiles ?* - contribue à inscrire la problématique de la référence dans une perspective de cognition située. Cet article traitera spécifiquement de la construction dynamique en discours des couleurs en tant qu'objets cognitifs, à travers l'étude des relations entre pratiques perceptives, lexicalisations en discours, et représentations. L'analyse de discours recueillis dans des situations perceptives variées nous permettra de mettre en évidence la construction en discours de différentes conceptualisations de la couleur à travers, notamment, l'identification de fonctionnements syntaxiques, indices de différents modes de construction de la référence.

2. Problématique

2.1. Cadre théorique

L'étude des pratiques langagières est envisagée comme un moyen d'accès privilégié pour comprendre les modalités de construction des représentations sensibles et des connaissances du monde perçu. Si les représentations sensibles sont spécifiques à chaque locuteur, les connaissances collectives résultent de facteurs historiques, sociaux, culturels et linguistiques. A travers l'expression discursive d'un locuteur, les représentations subjectives et les connaissances partagées par l'ensemble de sa communauté linguistique sont exprimées et co-construites. Ces hypothèses autorisent à émettre des inférences à propos des représentations et connaissances, à partir des analyses linguistiques (Dubois, 2000).

Les méthodes expérimentales classiques largement utilisées **en psychologie cognitive**, centrées sur l'étude analytique d'un processus spécifique, à partir d'une description du monde empruntée aux sciences de la nature, se placent dans une perspective dite « bottom-up », des bas niveaux de traitement aux représentations centrales, qui tendent à maximiser le rôle des contraintes physiologiques des récepteurs et du même coup à valider des hypothèses universalistes. En contrepoint, l'approche holistique de la cognition située qui est à la base de notre réflexion, prend également en compte les traitements descendants, considérant les

¹ dans le cadre d'un thèse CIFRE : Collaboration CNRS/PSA Peugeot Citroën/Univ. Paris 3.

constructions cognitives comme le résultat d'une expérience à la fois individuelle et partagée, et met davantage en évidence la diversité culturelle (Rastier, 1991).

En linguistique, la plupart des travaux sont structurés à partir de catégories linguistiques, telles que les catégories grammaticales, les formes linguistiques, les processus énonciatifs. Notre approche nous conduit à une problématique qui prend en compte ces différents plans linguistiques comme autant de plans de description permettant de comprendre comment langage et cognition co-construisent dynamiquement en discours significations et catégories cognitives à partir de l'expérience sensible et des ressources langagières disponibles chez les locuteurs.

2.2. Les couleurs

2.2.1. Débat entre relativisme et universalisme

La dénomination et la catégorisation des couleurs, terrain emblématique du débat entre universalisme et relativisme, ont suscité tant en psychologie qu'en linguistique, de nombreux travaux (pour une revue voir Taylor, 2002), à la suite de ceux de Berlin *et al.* (1969). Les théories relativistes actuelles en linguistique (Lucy, 1997) et en anthropologie (Foley, 1997) reprenant l'héritage relativiste (cf. par ex. Conklin, 1964), insistent sur le mode de questionnement utilisé dans les expériences pour la mise en évidence des concepts de couleur. En effet, dans les travaux issus de Berlin *et al.* (1969), les couleurs sont posées comme étant d'évidence adéquatement représentées par des *color chips* et par des *basic color terms*, le sens lexical s'inscrivant dans une sémantique à valeur référentielle.

De plus, malgré l'abondance d'études « universalistes » sur la vision des couleurs en neurosciences, certaines études (cf. Zeki, 2005 par ex.) mettent l'accent sur le fait que les contraintes physiologiques ne suffisent pas pour expliquer la vision des couleurs ni la catégorisation des couleurs. Les couleurs, en tant que concepts, sont des constructions de l'esprit (*mind*), et non du cerveau (*brain*), qui reposent non pas uniquement sur les propriétés des récepteurs sensoriels mais aussi sur le type d'activité dans laquelle les individus sont engagés. En psychologie cognitive, Roberson *et al.* (2005) apportent également des arguments empiriques en faveur d'une théorie psychologique relativiste, remettant en cause l'évidence du concept de couleur. Cependant l'ensemble de ces travaux posent les couleurs comme concepts historiquement situés.

2.2.2. Des objets culturels, linguistiques et cognitifs à réévaluer

Nous considérons ici les couleurs comme des **objets culturels situés**, à travers leur histoire (Pastoreau, 2000), la diversité des expériences et des pratiques de la couleur (Goodwin, 1997), ainsi que la façon dont les locuteurs négocient ces expériences au travers des contraintes linguistiques et des productions discursives. Les couleurs apparaissent alors comme **des concepts au statut non unifié**. Certaines cultures ne possèdent pas dans leur langue de termes de couleur, n'ayant pas conceptualisé la couleur comme une construction cognitive unifiant dans une « dimension » la diversité des expériences des couleurs. D'autres n'ont pas centré leur attention sur les dimensions de teinte, de saturation et de contraste qui caractérisent la description scientifique occidentale du « phénomène coloré », mais sur des propriétés telles que humide/sec, uni/rayé, naturel/teint (Wierzbicka, 1996).

Les limites du concept de « basic color term » sont également mises en évidence dans des travaux linguistiques sur les langues romanes. Au travers d'analyses lexicographiques, ainsi que de leur fonctionnement en discours, ces termes s'avèrent comme des *termes libres* en regard d'autres termes dépendant de leur contexte d'énonciation (Kristol, 1978). Enfin, l'analyse de la diversité des modes de dénominations de couleurs fournit des éléments de preuve de la diversité des représentations cognitives des couleurs : différents niveaux de

catégorisation, couleur comme dimension abstraite ou rattachée à un objet, etc. (Dubois *et al.*, 2003). À la lumière de ces différents travaux, les couleurs se révèlent **un objet cognitif à réévaluer en discours, à travers différentes expériences et pratiques perceptives situées.**

3. Les couleurs en discours

3.1. Démarche de l'enquête

La démarche globale de la thèse consiste à enregistrer des discours dans des situations perceptives contrastées. Parmi les différents discours recueillis auprès de locuteurs français à propos de leur expérience sensorielle et de leurs pratiques des habitacles automobiles, nous nous intéressons ici à des discours suscités dans des dispositifs expérimentaux (simulations visuelles sur un écran plasma® ou dans une salle de réalité virtuelle) et dans des habitacles réels de voiture. L'enquête a été réalisée avec N. Cavalier (2003).

3.2. Construction de l'expérience

Les locuteurs décrivent leur ressenti à un enquêteur, face à plusieurs habitacles variant selon le mode de présentation et selon les propriétés de couleur, de forme et de matière.

3.2.1. Matériel expérimental

Le premier mode de présentation (*dispositif 2D*) est un écran Plasma® proposant aux locuteurs des images statiques bidimensionnelles représentant l'avant de l'intérieur d'un véhicule. Le second (*dispositif 3D*) fait appel aux technologies de « réalité virtuelle ». Les images sont projetées sur quatre parois (latérales, frontale, au sol) d'une salle cubique (CAVE®). Le locuteur, assis dans un siège auto, porte des lunettes stéréoscopiques munies d'un capteur de mouvement, qui lui permettent de voir des images tridimensionnelles calculées en temps réel par rapport à sa position. Le dernier mode de présentation (*dispositif REEL*) consiste en deux véhicules à l'arrêt.

Les ambiances d'habitacle présentées dans les différents dispositifs varient selon la forme (deux architectures), la couleur (beige et noir) et la matière (bois *et* aluminium).

3.2.2. Locuteurs

15 locuteurs français, 8 femmes et 7 hommes âgés de 23 à 74 ans, participent à l'étude. Tous sont peu familiers des systèmes de réalité virtuelle, n'ayant pas de pratique, personnelle ou professionnelle, de ces technologies.

3.2.3. Guide d'entretien, protocole de recueil, constitution des corpus

Chaque locuteur est interviewé via la méthode des entretiens semi-directifs. La construction du guide d'entretien suit les principes de l'entretien semi-directif (Blanchet, 2003), allant du général au spécifique et privilégiant les questionnements ouverts (ex. : *Qu'est-ce qui, pour vous, est important dans cet intérieur ?*). Les propriétés sensorielles (couleurs, formes ...) ne sont pas mentionnées par l'intervieweur dans ses relances.

Les locuteurs voient successivement deux ambiances dans chaque dispositif (2D puis 3D puis REEL) et une dernière ambiance dans le dispositif 3D.

Les 15 entretiens sont enregistrés en audio sur Mini Disc et durent 90 à 150 minutes. Ils sont retranscrits intégralement. Trois corpus sont constitués, correspondant aux trois dispositifs de recueil : corpus 2D, 3D et REEL.

3.3. Analyses et Résultats

3.3.1. Hypothèses et catégories d’analyse

Notre principale hypothèse concerne les constructions cognitives impliquées dans l’expérience perceptive et résultant de cette expérience, les indices linguistiques en discours de ces phénomènes perceptifs et les inférences cognitives qu’ils permettent de réaliser.

Nous faisons l’hypothèse que les modes de présentation des couleurs (des intérieurs) influencent la perception et le jugement des locuteurs. En effet, différents types d’expérience sensible sont proposés :

- complète (dispositif REEL) *versus* partielle (dispositifs 2D et 3D),
- partagée par le locuteur et l’intervieweur (2D et REEL) *versus* individuelle (3D),
- familière (2D et REEL) *versus* non familière (3D).

Nous postulons que ces modes de présentation suscitent différents modes de référenciation, et nous proposons d’identifier et de caractériser ces différents modes de construction dynamique de la référence (cf. Mondada *et al.*, 1995) à travers la diversité des formes linguistiques en discours, en repérant des indices linguistiques pertinents. Au-delà du repérage lexical, notre hypothèse principale concerne les fonctions syntaxiques des termes de couleur. Il s’agit de vérifier si différentes conceptualisations de la couleur peuvent être inférées à partir du type d’inscription syntaxique des termes de couleur en discours. L’inscription en discours d’un terme de couleur comme substantif *versus* comme adjectif, construira et témoignera davantage d’une couleur respectivement comme entité du monde *versus* comme propriété. De plus, un adjectif de couleur, selon qu’il est épithète ou attribut d’un substantif, constituera soit l’indice d’une propriété rattachée à un objet de manière identitaire, soit l’indice d’une construction dynamique en discours de la référence à l’objet désigné par le substantif².

3.3.2. Résultats lexicaux

Une première analyse lexicale du corpus (pris dans sa globalité) permet de faire l’inventaire des termes de couleur employés et de montrer l’insuffisance de la seule prise en compte des 11 *basic color terms* de Berlin *et al.*, comme le montre la diversité des termes recueillis (cf. Tableau 1 ci-dessous). Sont distingués dans ce tableau les termes désignant la couleur de manière générique (*couleur, nuancé*), des termes spécifiques de couleur (*beige, gris, écru* ...).

Termes de couleur (37 types / 890 occurrences)	
Génériques (15 types / 379 occurrences)	Spécifiques (22 types / 511 occurrences)
<i>couleur</i> 292, <i>coloris</i> 20, <i>teinte</i> 16, <i>contraste</i> 13, <i>ton</i> 11, <i>uni</i> 8, <i>dégradé</i> 3, <i>nuance</i> 3, <i>tonalité</i> 3, <i>transparent</i> 3, <i>coloration</i> 2, <i>monochrome</i> 2, <i>bicolore</i> 1, <i>contrasté</i> 1, <i>nuancé</i> 1, <i>teinté</i> 1	<i>beige</i> 117, <i>gris</i> 108, <i>noir</i> 98, <i>bleu</i> 72, <i>blanc</i> 34, <i>marron</i> 19, <i>vert</i> 17, <i>écru</i> 16, <i>rouge</i> 6, <i>orange</i> 4, <i>ivoire</i> 3, <i>crème</i> 2, <i>jaune</i> 2, <i>noircir</i> 2, <i>orangé</i> 2, <i>rosé</i> 2, <i>turquoise</i> 2, <i>blond</i> 1, <i>grisé</i> 1, <i>marronâtre</i> 1, <i>violet</i> 1

Tableau 1 - Termes de couleur dans le corpus total

Les termes génériques représentent 42,6% des occurrences totales de termes de couleur et les termes spécifiques 57,4%. Parmi eux, on dénombre 10 « termes de base » (361 occ.) représentant 40,6% du lexique. D’autres termes considérés comme spécifiques d’un champ lexical (*beige*), issus d’un objet source (*ivoire*) ou construits sur un autre terme de couleur³ (*marronâtre*) sont également recensés (149 occ. soit 16,7%).

Observant dans ces discours sensibles une diversité des lexicalisations de la couleur, une analyse de l’inscription en discours de trois termes spécifiques, les plus cités (36,3% des

² Sur le statut cognitif de l’adjectif, voir notamment Honeste (2004).

³ Ces trois critères (non monolexémie, référence à une source et spécificité lexicale) excluent, pour Berlin *et al.*, la possibilité de considérer ces termes comme basiques.

occ.), est proposée pour appréhender le statut cognitif et les conceptualisations de la couleur qu'ils désignent et construisent en discours.

3.3.3. Inscription en discours de beige, gris et noir

Le tableau 2 présente des exemples d'inscription en discours de *beige*, *gris* et *noir*.

Exemples extraits du corpus	
beige	- c'est un beige un espèce de beige un beige rosé - c'est vrai que p(eu)t-être le beige là en l'occurrence i(l) fait penser à du cuir
gris	- sous l(e) pare-brise là tout est gris bon y a du beige quoi mais - ça m'a l'air d'être ça m'a l'air euh aluminium brossé j(e) le vois j(e) le vois gris clair
noir	- i(l) y a cette espèce de tissu noir qui doit être un peu qui doit un peu amortir - mais les rétro sont noirs et ça c'est bien parce qu'au moins on n'est pas attiré par le rétro

Tableau 2 - Exemples extraits du corpus

a. Catégories syntaxiques

Les fonctions syntaxiques occupées par ces 3 termes de couleur en discours sont présentées dans le tableau 3 à travers la répartition entre substantif et adjectif, exprimée en pourcentage d'occurrences pour chacun des trois termes.

Terme	substantif	adjectif
beige	58,1%	41,9%
gris	50%	50%
noir	36,4%	73,6%

Tableau 3 - *beige*, *gris* et *noir* substantifs ou adjectifs

Les trois termes de couleur n'occupent pas les mêmes fonctions syntaxiques. Lorsque *beige* apparaît davantage comme substantif, *noir* est très majoritairement employé comme adjectif et *gris* apparaît de manière équivalente dans les deux fonctions syntaxiques.

b. Substantifs

En tant que substantifs, les trois formes ne s'inscrivent pas dans les mêmes constructions syntaxiques nominales. Ainsi *beige* apparaît davantage comme forme simple det+beige (*le beige*) que *noir* ou *gris* (44,9% des occurrences contre 35,2% et 37,5%). De plus, *noir* apparaît rarement inscrit dans une forme de type det+noir+adj (4,8%) à la différence de *beige* et *gris* (14,5 et 14,8% des occurrences).

Les formes nominales et adjectivales associées diffèrent également selon la couleur désignée. Si les trois formes sont associées à des noms désignant des éléments de l'habitable (*le beige du siège*, *le gris de la colonne centrale*, *le noir du levier*), *beige* est plus régulièrement associé à d'autres noms de couleur (*une gamme de beige marron*). *Noir* et *gris* sont eux plus souvent associés à des noms de matière (*un noir sur un tissu*, *le gris du plastique*). Enfin, *gris* est souvent associé à des adjectifs renvoyant aux propriétés d'intensité de la nuance (*un gris très très clair*).

c. Adjectifs

En tant qu'adjectifs, les trois termes ont également des modes d'inscription différents :

Adjectif	attribut	épithète
beige	36,7%	73,3%
gris	72,2%	37,8%
noir	53,6%	46,4%

Tableau 4 - *beige*, *gris* et *noir*, attributs ou épithètes

Beige apparaît principalement comme épithète, *gris* comme attribut et *noir* dans les deux fonctions.

Parmi les formes nominales et adjectivales associées, on repère des noms désignant des objets que viennent qualifier principalement les épithètes *gris* et *beige* (*les sièges beiges, la portière grise*). On dénombre également des noms de matière, soit qualifiés par *noir* épithète (*tissu noir, un plastique noir noir*), soit spécifiant l'adjectif de couleur (*gris métal, gris plastoc, beige imitation cuir*). Enfin, si *gris* et *noir* sont parfois précédés d'adverbes (*trop, hyper, plutôt, presque*), aucun adverbe ne vient modaliser l'adjectif *beige*.

d. Inférences cognitives

Beige, majoritairement employé comme substantif simple précédé d'un déterminant défini (*le beige*) ou accompagné d'un adjectif épithète, contribue à construire la couleur ainsi désignée comme une entité du monde que les locuteurs peuvent évaluer. Les formes complexes nominales construites autour du substantif *beige* construisent dès lors une entité pouvant être qualifiée (*les beige très clairs*) ; spécifiée par un objet (*le beige de la portière*), une matière source (*le beige [...] plutôt cuir*) ; et évaluée (*le beige plutôt luxueux*).

Gris employé comme substantif et comme adjectif attribut renvoie à deux statuts très différents. Sous forme substantivale, il renvoie au gris comme entité du monde plus (*un gris*) ou moins autonome (*le gris du plastique*). En tant qu'adjectif, il renvoie à une propriété d'un objet (*la voiture grise*) ou d'une couleur (*des tons gris*). Mais, très majoritairement attribut, il contribue à construire la référence en discours (*bah là la radio c'est toujours gris*).

Noir employé majoritairement comme adjectif se fait l'indice d'une couleur comme propriété d'une matière ou d'un élément (*le plastique noir noir, un bouton noir*).

Le contraste entre l'adjectif *beige* presque toujours épithète et les adjectifs *gris* et *noir* le plus souvent attributs et modalisés par différents adverbes contribue également à différencier les statuts cognitifs des couleurs désignées. Alors que *beige* réfère à une propriété inhérente à l'objet qualifié, *gris* et *noir* renvoient à une dimension « quantifiable » (*hyper gris...*).

Enfin, on remarque également le contraste entre la couleur comme propriété de la matière désignée par les constructions de type Nmatière+Adjcouleur (*le tissu noir*) et la couleur comme indice d'une matière avec des constructions de type Ncouleur+Nmatière (*le gris métal, le beige plutôt cuir*) qui concernent davantage *gris* et *beige*.

3.3.4. Inscription en discours selon les modes de présentation

a. Répartition des termes de couleur

Considérant l'ensemble des occurrences de termes de couleur relevées dans les discours produits devant chaque mode de présentation, quelques différences apparaissent. Les dispositifs 2D, 3D et REEL ont suscité respectivement 55, 70 et 60% d'occurrences de termes spécifiques de couleur.

Plus spécifiquement, le Tableau 5 présente la répartition des occurrences de *beige*, *gris* et *noir* dans les 3 corpus.

Terme	CORPUS 2D	CORPUS 3D	CORPUS REEL
beige	42 occ.	35 occ.	40 occ.
gris	59 occ.	27 occ.	22 occ.
noir	18 occ.	15 occ.	65 occ.

Tableau 5 - Occurrences de *beige*, *gris* et *noir* dans les 3 corpus

Beige présente un nombre quasi identique d'occurrences dans les trois corpus alors que *gris* et *noir* ont des répartitions plus hétérogènes et nettement opposées. *Gris*, le plus cité dans le corpus 2D, y est 3 fois plus présent que *noir*. A l'inverse, dans le corpus REEL, il est le moins employé, 3 fois moins que *noir*, terme spécifique le plus cité.

Le tableau 6 propose quelques extraits des différents corpus :

Exemples extraits du corpus	
2D	- sous l(e) pare-brise là tout est gris bon i(l) y a du beige quoi mais euh ouais - euh ça ça (n')a rien ni de joli pas joli quoi hein c'est du gris
3D	- j'ai l'impression que ça fait plus cuir la partie beige que plastique - euh j(e) vois gris j(e) sais pas si c'est noir hein
REEL	- (...) mais alors c'est bizarre parce que le beige du tableau d(e) bord n'est pas le même que le beige de la portière - (...) on a du beige et noir et là c'est un gris foncé et et proche du noir

Tableau 6 - Exemples extraits des corpus 2D, 3D et REEL

b. Positions syntaxiques

Les positions syntaxiques occupées sont plus ou moins stables. *Beige* majoritairement substantif dans les trois corpus, l'est davantage dans le corpus 3D (62,9%). *Gris* est majoritairement adjectif dans le corpus 3D (63%), et davantage substantif (55,9%) dans le corpus 2D. *Noir*, toujours davantage présent comme adjectif, l'est majoritairement en 2D et 3D (66,7 et 73,3%).

Observant les positions respectives d'attribut et d'épithète des adjectifs *beige*, *gris* et *noir*, des distinctions s'opèrent également. *Beige* est toujours employé majoritairement comme épithète, *gris* toujours comme attribut. *Noir*, plus épithète dans les corpus 2D et REEL est quasi exclusivement attribut dans le corpus 3D (90,9%).

c. Inférences cognitives

Beige, présent dans les trois corpus comme substantif, renvoie à **une couleur stable, consensuelle** que les locuteurs peuvent évaluer. En tant qu'adjectif, il est presque toujours épithète, rattaché à un nom qu'il qualifie sans équivoque. En revanche, *gris* et *noir* se font l'indice d'une **construction problématique de la couleur en discours** notamment dans le corpus 3D (ex : (...) *et euh j(e) vois gris j(e) sais pas si c'est noir hein*). En effet, si les locuteurs emploient très majoritairement *gris* dans le corpus 2D, ils expriment leur incertitude devant les ambiances 3D (c'est noir ou gris ?) pour ensuite re-qualifier la couleur comme *noir* dans les véhicules réels. Cette incertitude est en partie liée à la différence de luminosité des dispositifs 2D et 3D par rapport aux véhicules réels.

Le corpus 3D suscite également plus d'emplois de *beige*, *noir* et *gris* en tant qu'attributs. **Devant un dispositif non familier**, dont en outre ils ne partagent pas la perception avec leur intervieweur (eux seuls portent les lunettes 3D), **les locuteurs doivent construire la référence** et utilisent comme moyen les adjectifs de couleur qui désignent **des propriétés construisant la matérialité** des objets représentés visuellement à partir de la couleur. A l'inverse, dans le dispositif réel, n'ayant plus besoin de construire la référence, familière et partagée avec leur interlocuteur, ils peuvent **évaluer les couleurs en tant qu'entités du monde**, ce que nous indiquerai^{ent} entre autres les proportions plus élevées de substantifs pour *noir* et *gris*.

4. Conclusions et Perspectives

L'analyse de l'inscription syntaxique en discours de trois termes de couleur nous permet d'identifier différents plans de construction des couleurs tant dans leur fonctionnement en discours qu'en cognition et nous confirment l'importance du rôle des formes linguistiques dans la construction des concepts. Les statuts cognitifs inférés sont riches et diversifiés : les couleurs sont à la fois inscrites en discours comme entité du monde (*le beige*), comme propriété (*plastique noir*) ou comme indice d'un objet, d'une matière (*ça fait plus cuir la partie beige que plastique*). Elles contribuent également à la construction de la référence notamment dans des scènes perceptives visuelles tridimensionnelles en tant que reproductions partielles et non familières du monde sensible. Le contraste entre *beige*, *gris* et *noir* met en évidence la non-correspondance de statut cognitif pour différentes couleurs. Au

contraire, les pratiques culturelles qui sous-tendent chaque couleur et la pratique spécifique mise en place lors du recueil des données amènent à des couleurs plus ou moins différenciées cognitivement selon les couleurs et les modes de présentation « expérimentés » par les locuteurs.

La méthodologie mise en œuvre pour étudier les couleurs en discours permet de faire un inventaire des ressources linguistiques pour ensuite étayer nos hypothèses sur les relations entre discours, perception et cognition. Cette approche nous amène à considérer une diversité des constructions cognitives à propos des phénomènes sensibles. On retrouve ici, au sein de la langue et de la culture française, la diversité des modes d'inscription des phénomènes colorés observés dans la diversité des cultures.

Références bibliographiques

- BERLIN, Brent & KAY, Paul (1969). *Basic color terms : their universality and evolution*. Berkeley/Los Angeles : Univ. of California Press.
- BLANCHET, Alain (2003). *Dire et faire dire : l'entretien*. Paris : Armand Colin.
- CAVELIER, Noémie (2003). *Descriptions d'ambiances réelles et virtuelles : Analyses linguistique et cognitive de différents dispositifs de représentation d'un habitacle automobile*. DEA de Sciences du Langage, Université de Paris 3.
- CONKLIN, Harold (1964). Hanunoo color categories, in *Language in culture & society : a reader in linguistic and anthropology*, Hymes, Dell (Ed). N.Y : Harper & Row, 189-192.
- DUBOIS, Danièle (2000). Categories as acts of meaning : the case of categories in olfaction and audition, *Cognitive Science Quarterly*, 1: 35-48.
- DUBOIS, Danièle & GRINEVALD, Colette (2003). En voir de toutes les couleurs : processus de dénomination des couleurs et constructions cognitives, in *Langues et cognition*, Vandeloise, C. (Ed). Paris : Lavoisier, 79-113.
- FOLEY, William (1997). *Anthropological linguistics*. Malden/Blackwell : Cambridge Univ. Press.
- HONESTE, Marie-Luce (2004). Approche cognitive de la fonction adjectivale, in *L'Adjectif en français et à travers les langues*, François, J. (Dir). Caen : PUC, p.135-150.
- GOODWIN, Charles (1997). The Blackness of Black: Color Categories as Situated Practice, in *Discourse, Tools and Reasoning: Essays on Situated Cognition*, Resnick, Lauren, Säljö, Roger, et al. (Eds). New York : Springer-Verlag, 111-40.
- KRISTOL, Andres (1978). *Color. Les langues romanes devant le phénomène de la couleur*. Berne : Francke.
- LUCY, John (1997). The linguistics of "color", in *Color Categories in Thought and Language*, Hardin, C.L. & Maffi, L. (Eds). Cambridge: Cambridge Univ., 320-346.
- MONDADA, Lorenza & DUBOIS, Danièle (1995). Construction des objets de discours et catégorisation : une approche des processus de référenciation, *Tranel, Travaux Neuchâtelois de Linguistique*, 23, 273-302.
- PASTOUREAU, Michel (2000). *Bleu, histoire d'une couleur*. Paris : Seuil.
- ROBERSON, Debi, DAVIDOFF, Jules, et al. (2005). Color categories in Himba: Evidence for the cultural relativity hypothesis, *Cognitive Psychology*, 50: 378-411.
- RASTIER, François (1991). *Sémantique et recherches cognitives*. Paris : PUF.
- TAYLOR, John (2002). *Cognitive grammar*. Oxford: Oxford University Press.
- WIERZBICKA, Ana (1996). *Semantics : primes and universals*. Oxford : Oxford Univ. Press.
- ZEKI, Semir (2005). *The construction of colour in the brain*. Colloque IRIS, Essence et Sens de la Couleur, Nancy 09/11/05.

ÉVALUATION DE LA REPRÉSENTATION INTRA- ET INTER-HÉMISPHERIQUE DU LANGAGE CHEZ LES SUJETS SAINS : ÉTUDE COMPORTEMENTALE EN CHAMP VISUEL DIVISÉ ET ÉTUDE IRMf¹

Émilie COUSIN

emilie.cousin@upmf-grenoble.fr

Université Pierre Mendès France

CNRS UMR 5105

BP 47, 38040 Grenoble cedex 9

Abstract : fMRI is a non-invasive means of assessing hemispheric language dominance (inter-hemispheric representation) and areas involved in phonology and semantics (intra-hemispheric representation). However, it still presents one important limitation: the performed task in fMRI. The aim of the present behavioural experiment is to explore the degree of dominance of several language tasks frequently used in fMRI paradigms.

1. Introduction

1.1. *Appréhension à écrire*

Le cerveau est constitué des deux hémisphères cérébraux, un droit et un gauche, reliés par plusieurs faisceaux de substance blanche. Ces faisceaux permettent un traitement rapide de l'information par les deux hémisphères induisant ainsi la production d'une pensée intégrée. Beaucoup de faits accumulés montrent que les deux hémisphères ne sont pas identiques, ni dans leur configuration anatomique, ni dans leurs aptitudes et leur organisation fonctionnelle. En effet, on parle de dominance hémisphérique ou de représentation inter-hémisphérique. Ce terme est défini par le fait que chacun des hémisphères est en fait spécialisé pour certaines fonctions propres. Par exemple, concernant le langage, 90% d'entre nous utilisent préférentiellement l'hémisphère gauche: on parle alors d'hémisphère dominant.

1.2. *Représentation intra-hémisphérique*

Au sein même de cet hémisphère dominant, des régions anatomiques particulières assurent différents aspects du langage. Ainsi, certaines régions prennent en charge les aspects phonologiques (*i.e.* transformation graphème/phonème) du langage écrit tandis que d'autres sont impliquées dans les aspects sémantiques du langage (*i.e.* accès au sens). On parle cette fois de représentation intra-hémisphérique. L'évaluation de la représentation inter- et intra-hémisphérique est requise dans le cas de patient présentant une épilepsie pouvant être traitée chirurgicalement par excision de la zone cérébrale responsable de la pathologie. Cette zone est malheureusement très souvent placée à proximité d'une région prenant en charge la fonction langagière. C'est pourquoi un bilan pré-chirurgical d'évaluation de la représentation inter- et intra-hémisphérique est nécessaire afin d'éviter les impairs dus à la chirurgie et plus particulièrement une aphasie (*i.e.* trouble du langage affectant l'expression ou la compréhension du langage parlé ou écrit causé par une lésion cérébrale).

¹ IRMf : Image par Résonance Magnétique fonctionnelle

1.3. Évaluation de la représentation inter- et intra-hémisphérique

1.3.1. Méthode invasive: le test de Wada

Une méthode invasive d'évaluation de la représentation inter- et intra-hémisphérique du langage chez des patients est le test de Wada mis au point par Juhn Wada en 1958. Cette procédure consiste à injecter un barbiturique dans une des carotides internes. Ce qui a pour effet d'anesthésier l'hémisphère du côté opposé à l'injection. Lorsque l'anesthésie hémisphérique est installée, le médecin peut alors évaluer les éventuels troubles du langage, d'une part en observant la manière dont le patient continue (ou non) à compter et, d'autre part, en lui faisant passer des tests de langage supplémentaires. Bien que ce test soit considéré comme étant la méthode de référence («clinical gold standard method»), son caractère invasif peut entraîner de telles difficultés d'interprétation (le patient parle toujours après anesthésie de l'hémisphère droit mais aussi après anesthésie de l'hémisphère gauche: le test n'est donc pas conclusif !) et de telles complications potentielles que la seule solution serait une méthode non invasive alternative. Il faut également ajouter à ces complications le fait que le test de Wada permet uniquement de conclure sur la représentation inter-hémisphérique (*i.e.* dominance d'un hémisphère) rendant impossible une évaluation de la représentation intra-hémisphérique (*i.e.* cartographier les zones précises qui prennent en charge la fonction langagière).

1.3.2. Une méthode non invasive alternative: l'Imagerie par résonance magnétique fonctionnelle

L'imagerie par résonance magnétique fonctionnelle pourrait, grâce à son caractère non invasif, constituer une méthode alternative au test de Wada. En effet, des corrélations très fortes ont été mises en évidence entre les résultats fournis par ces deux méthodes. L'IRMf a été appliquée chez des patients ayant déjà subi le test de Wada pour comparer les dominances hémisphériques obtenues par ces deux méthodes (*Baciu et al., 2001*). Les résultats obtenus entre la dominance évaluée avec le test de Wada et celle évaluée avec l'IRMf sont assez concordants d'une manière générale. De plus, l'IRMf présente l'intérêt de permettre d'inférer les capacités langagières des patients en post-chirurgical (par la localisation précise des aires impliquées dans le langage). Ceci étant, l'IRMf n'indique pas toujours correctement l'hémisphère dominant. L'opinion la plus répandue est que, en l'état actuel des recherches, l'IRMf ne peut pas remplacer le test de Wada dans la pratique clinique (*Baciu et al., 2003*). On peut attribuer à ces difficultés et à ce léger désaccord entre les résultats obtenus par le test de Wada et par la technique IRMf une cause principale: le choix des tâches de langage réalisées par le sujet dans l'IRM.

1.3.3. Problème des tâches utilisées dans l'IRMf

Les tâches utilisées pour des expérimentations IRMf concernant la détermination de la dominance hémisphérique doivent répondre à plusieurs critères. Le plus important est la nécessité que ces tâches induisent une activation soutenue dans l'hémisphère dominant. D'autre part, ces tâches doivent permettre de mettre en évidence les aspects sémantique et phonologique du langage (représentation intra-hémisphérique) et être facilement réalisable pour le sujet. Au vu de ces critères, il n'existe pas de véritable consensus concernant le choix des tâches pour l'évaluation de la représentation inter- et intra-hémisphérique.

1.3.4. Méthode comportementale: la technique en champ visuel divisé

La technique en champ visuel divisé est une méthode comportementale qui permet d'approcher la latéralisation fonctionnelle du cerveau normal. Elle se base sur les propriétés

des voies visuelles qui sont croisées et implique la présentation de stimuli dans un hémichamp visuel (hémichamp visuel droit vs hémichamp visuel gauche). Les stimuli présentés d'un côté d'un point de fixation central seront ainsi projetés en premier lieu à l'hémisphère controlatéral : un mot présenté dans l'hémichamp visuel droit (CVd) sera présenté en premier lieu sur le cortex visuel de l'hémisphère gauche (HG) tandis qu'un mot présenté dans l'hémichamp visuel gauche (CVg) sera présenté au cortex visuel de l'hémisphère droit (HD). L'efficacité des deux systèmes CVg-HD et CVd-HG peut être comparée lors de tâches très variées et permet ainsi de faire des inférences sur la spécialisation fonctionnelle des hémisphères (en fonction du matériel présenté) (*Hellige, 1996*). En effet, le postulat de base stipule que la performance sera meilleure si un stimulus a été présenté initialement à l'hémisphère spécialisé pour traiter le matériel (et/ou réaliser la tâche) (*Springer & Deutsch, 2000*). Ce postulat s'applique également aux temps de réponses qui seront plus courts.

1.4. Objectifs de l'étude

Notre but est de pallier les limites de l'IRMf citées précédemment concernant le choix de la tâche. Cette étude comportementale vise tout d'abord à choisir une tâche de langage induisant une latéralisation prégnante chez les sujets sains dans le but ultérieur d'investiguer leur représentation inter- et intra-hémisphérique dans une étude IRMf. En outre, nous voulons également étudier deux aspects essentiels du langage lors de l'examen IRMf : la phonologie et la sémantique. Dans ce but, nous avons réalisé deux épreuves en champ visuel divisé, l'une dite «phonologique», composée de deux tâches en champ visuel divisé à prédominance phonologique, et l'autre dite «sémantique», composée de deux tâches en champ visuel divisé à prédominance sémantique. Les résultats de ces épreuves nous permettront d'arrêter notre choix sur une tâche phonologique et sur une tâche sémantique puisque nous pourrions déterminer quelles sont les tâches les plus latéralisées parmi les deux tâches phonologiques et parmi les deux tâches sémantiques, respectivement. Nous avons choisi d'évaluer 4 tâches : deux tâches phonologiques fréquemment utilisées en IRMf : la détection de rime et la détection de phonème et deux tâches sémantiques également fréquemment utilisées : la catégorisation «vivant» et «comestible». Ces tâches seront ensuite utilisées dans une étude IRMf dans l'évaluation de la représentation inter-hémisphérique mais également intra-hémisphérique (*i.e.* phonologie et sémantique).

2. Étude comportementale exploratoire en champ visuel divisé

2.1. Méthode

2.1.1. Participants

Le choix des participants a été déterminé par leur latéralité manuelle et leur sexe. En effet, les sujets droitiers et gauchers présentent des différences concernant la spécialisation hémisphérique. En effet, de nombreuses recherches s'accordent sur une latéralisation moindre des gauchers (*Eviatar et al., 1997*). De même, il existe un large débat dans la littérature concernant les différences sexuelles. Une latéralisation plus prégnante chez les hommes que chez les femmes a été montrée grâce à une étude menée en IRM fonctionnelle (*Shaywitz et al., 1995*). Trente huit hommes âgés de 19 à 57 ans (âge moyen de 26,7 ans) ont donc participé à l'expérience. La latéralité manuelle de chaque participant a été investiguée grâce au questionnaire de latéralité d'Edinburgh (*Oldfield, 1971*). Tous les participants sont droitiers (Indices de latéralisation d'Edinburgh compris entre 0,8 et 1) et naïfs concernant le paradigme expérimental utilisé.

2.1.2. Matériels

a. Dispositif expérimental

Les sujets ont été testés individuellement. Chaque sujet était installé dans une pièce sombre face à un écran d'ordinateur (19" TM Ultra Scan P790 de résolution 1024 par 768 pixels). L'expérience a été réalisée à l'aide du logiciel E-Prime.

b. Stimuli

Epreuve phonologique

Deux listes de 96 pseudo-mots de 6 lettres ont été élaborées pour la construction de cette épreuve. Dans la première liste, la moitié des items contient le phonème /b/ à détecter. Sa position au sein de la séquence de lettres est contrebalancée afin que celui-ci puisse être identifié un nombre de fois équivalent aux 3 positions (*i.e.* début, milieu et fin du pseudo-mot). La seconde liste est également constituée d'une moitié d'items tests contenant la rime /é/. La fréquence des trois formes orthographiques (é, er & ée) de cette rime est contrebalancée afin de neutraliser un éventuel biais lié à la représentation visuelle orthographique. Des items contrôles ont également été élaborés. Ils sont constitués par un rectangle rouge de la taille du mot le plus long apparaissant 96 fois au cours de l'épreuve. Ces items contrôles ont été élaborés pour que les participants réalisent une tâche visuelle de bas niveau.

Epreuve sémantique

Deux listes de 96 mots de 6 lettres ont été élaborées pour la construction de cette épreuve. Dans la première liste, la moitié des items appartient à la catégorie sémantique «vivant» (plantes et animaux en excluant les aliments pour lesquels une violation de la dichotomie vivant/non vivant est constatée en neuropsychologie (*Capitani et al., 2003*) et l'autre moitié appartient à la catégorie sémantique «non-vivant». La seconde liste est également constituée d'une moitié d'items tests appartenant à la catégorie sémantique «comestible» (ingrédients *et* aliments) et l'autre moitié appartient à la catégorie sémantique «non-comestible». Pour chaque tâche au sein de chaque épreuve, 21 items du même type ont été élaborés (7 items corrects, 7 items incorrects et 7 items contrôles constitués par un rectangle rouge) pour la phase préalable d'apprentissage. Une police *courrier new* taille 24 a été utilisée pour chacun des items des deux épreuves.

2.1.3. Procédure expérimentale

Les consignes orales pour chaque tâche sont fournies au participant. Le sujet a ensuite la consigne écrite pour chacune des tâches avant et après la phase d'apprentissage. Le participant a pour consigne d'appuyer sur la clef de réponse soit lorsque la réponse à la tâche demandée est correcte soit lorsqu'il voit un item contrôle sinon il lui est demandé de ne rien faire. Ainsi, pour l'épreuve phonologique, il est demandé aux participants de détecter le phonème /b/ (tâche de détection de phonème) et de détecter la rime /é/ (tâche de détection de rime). Lors de l'épreuve sémantique, les participants sont assignés à appuyer sur le boîtier de réponses lorsque les mots appartiennent à la catégorie sémantique «vivant» (tâche de catégorisation «vivant») et «comestible» (tâche de catégorisation «comestible»). Tous les participants réalisent les deux épreuves (*i.e.* 2 tâches phonologiques et 2 tâches sémantiques) Pour chaque tâche, la présentation des items se fait comme suit. Un essai débute par la présentation centrale d'un point de fixation durant 500 ms afin de diriger le regard du participant au centre de l'écran. Il est immédiatement suivi de la présentation latéralisée du stimulus (champ visuel gauche ou champ visuel droit). Le temps de présentation du stimulus devant être assez court (inférieur à 150 ms afin d'éviter les saccades oculaires remettant le sujet en vision centrale), nous avons choisi 130 ms. L'item est suivi d'un masque (suite

d'étoiles de la taille du mot le plus long, ceci dans le but d'éviter les phénomènes de persistance rétinienne) durant 30 ms. L'ensemble de ces éléments est présenté sur un fond blanc et la présentation est aléatoire. Le nombre de réponses correctes est enregistré pour chaque participant. La variable de traitement utilisée lors des analyses statistiques est le pourcentage de bonnes réponses (%BR). Nous ne présentons pas les résultats en termes de temps de réponse du fait de la procédure «go/nogo» qui ne permet le recueil des temps de réponse que pour la condition «go».

2.1.4. Prédiction

Trois hypothèses peuvent être posées pour chaque épreuve:

1. Pour chaque tâche, nous nous attendons à un avantage du champ visuel droit (hémisphère gauche) sur le champ visuel gauche, c'est-à-dire à ce que le %BR soit significativement plus élevé dans le champ visuel droit que dans le champ visuel gauche.
2. Si une tâche est plus latéralisée que l'autre au sein de chaque épreuve, la différence de %BR entre les champs visuels devrait être plus importante pour cette tâche que pour l'autre.
3. Pour la tâche visuelle de bas niveau, nous nous attendons à aucune différence de %BR entre les deux champs visuels.

2.1.5. Résultats

Une ANOVA (analyse de variance) par participant a été conduite pour chacune des épreuves sur les pourcentages de réponses correctes. Le seuil de significativité α de rejet de l'hypothèse nulle a été fixé à .05 pour l'ensemble des tests statistiques présentés (logiciel STATISTICA 5.1).

a. Effet du champ visuel : tâche contrôle de bas niveau

Il n'y a pas d'effet principal du champ visuel de présentation pour la tâche de bas niveau ni pour l'épreuve phonologique ni pour l'épreuve sémantique.

b. Effet principal du champ visuel : Epreuves phonologique et sémantique

Pour l'épreuve phonologique, un effet significatif du champ visuel de présentation est observé avec un pourcentage de réponses correctes moyen plus élevé pour le champ visuel droit/hémisphère gauche (77%) (chpVD/HG) que pour le champ visuel gauche/hémisphère droit (chpVG/HD) (65%). Pour l'épreuve sémantique, un effet significatif du champ visuel de présentation est également observé avec un taux de réponses correctes moyen plus élevé pour le champ visuel droit/hémisphère gauche (71%) (chpVD/HG) que pour le champ visuel gauche/hémisphère droit (chVG/HD) (58%).

c. Effet de la tâche: Epreuves phonologique et sémantique

Il n'y a pas d'effet principal de la tâche ni pour l'épreuve phonologique ni pour l'épreuve sémantique, montrant ainsi que toutes ces tâches sont de difficulté équivalente.

d. Interaction entre tâches et champs visuels

L'interaction tâche * hémisphère va nous permettre de savoir si une des deux tâches active plus l'hémisphère dominant dans chaque épreuve. Elle se révèle significative pour l'épreuve phonologique tout comme pour l'épreuve sémantique signifiant ainsi que l'avantage du chp VD/HG est plus important (i) dans la tâche de détection de rimes que dans la tâche de détection de phonèmes et (ii) dans la tâche de catégorisation «vivant» que dans la tâche de catégorisation «comestible».

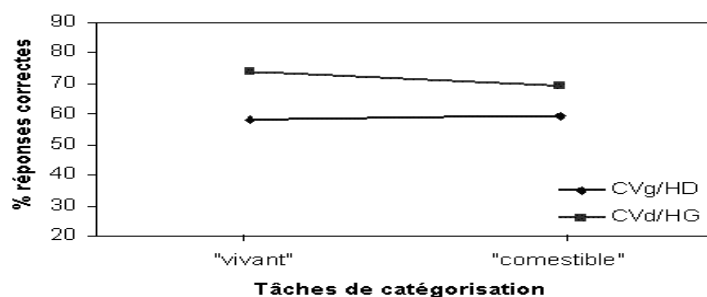


Figure 1 - Graphique d'interaction tâches * champ visuel de présentation pour l'épreuve phonologique.

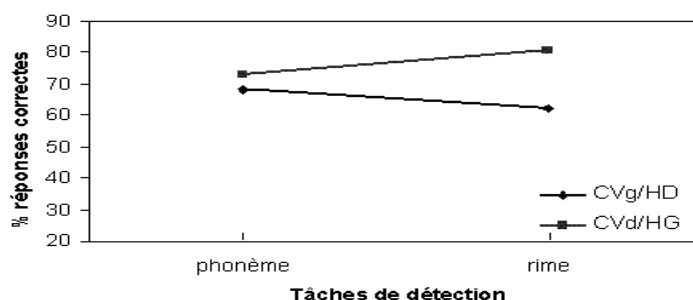


Figure 2 - Graphique d'interaction tâches * champ visuel de présentation pour l'épreuve sémantique.

2.1.6. Synthèse des résultats

Par rapport à nos hypothèses opérationnelles, nous mettons en évidence dans toutes les épreuves (*i.e.* phonologique et sémantique) et pour chacune des tâches de ces épreuves un avantage du champ visuel droit (hémisphère gauche) sur les performances montré par un effet significatif du champ visuel de présentation. De plus, dans chaque épreuves, l'interaction tâche * champ visuel de présentation est significative dans le sens où la tâche de détection de rime est plus latéralisée que la tâche de détection de phonème et que la tâche de catégorisation «vivant» est plus latéralisée que la tâche de catégorisation «comestible». De plus, l'analyse des performances pour les items contrôles montre que l'efficacité des 2 systèmes CVg/HD et CVd/HG est la même.

2.2. Discussion: étude comportementale exploratoire

En premier lieu, il est important de remarquer que, pour les deux épreuves sémantiques et phonologiques, nous avons observé une supériorité du champ visuel droit confirmant la dominance de l'hémisphère gauche pour des tâches de langage (que celles-ci soient à prédominance phonologique ou sémantique). Par ailleurs, nous n'avons observé aucune différence de performance entre les deux tâches suggérant que toutes les tâches sont de difficultés équivalentes. Enfin, aucune différence de performance n'a été observée pour la tâche contrôle visuelle de bas niveau. Ces éléments permettent de confirmer la validité de la procédure en champ visuel divisé.

2.2.1. Épreuve phonologique

Concernant l'épreuve phonologique, nous avons mis en évidence une latéralisation plus importante de l'hémisphère gauche dans la tâche de détection de rime que dans la tâche de détection de phonème. Ce résultat est concordant avec les données de la littérature montrant

une latéralisation prégnante de la tâche détection de rime (*Baciu et al., 2001*). La latéralisation moindre de la tâche de détection de phonème peut être expliquée par le fait que cette tâche s'apparente plus à une détection de lettre. En effet, il s'agit d'une limite de notre étude. Nous avons utilisé le phonème /b/. Or, ce phonème est grapho-phonétiquement transparent (*i.e.* la transformation de l'orthographe à la phonologie est régulière). Une équipe de recherche a suggéré que le niveau de transparence d'un phonème aurait un impact sur les performances puisque le traitement de phonèmes transparents semble requérir la participation des deux hémisphères (*Tremblay et al., 2004*). Par ailleurs, une étude en TEP² montre que, lors d'une détection de lettre, les activations des zones dédiées au langage sont réduites. En effet, les activations les plus prégnantes observées dans une telle tâche sont bilatérales et concernent les aires striées et extrastriées. Ces activations correspondraient aux traitements des formes visuelles et plus spécifiquement aux caractéristiques visuelles de la lettre à détecter (*Cappa et al., 1998*). Ainsi, ces données de neuroimagerie corroboreraient une latéralisation moindre de notre tâche de détection de phonème du fait de sa très forte composante visuelle qui pourrait être traitée par les deux hémisphères. Ces affirmations expliqueraient que la différence de performance entre les champs visuels de présentation soit réduite pour cette tâche.

2.2.2. Épreuve sémantique

Concernant l'épreuve sémantique, nous avons mis en évidence une latéralisation plus importante de l'hémisphère gauche dans la tâche de catégorisation «vivant» que dans la tâche de catégorisation «comestible». Ce résultat est concordant avec les données de la littérature en IRMf (*i.e.* montrant une activation robuste de l'hémisphère gauche ; (*Seghier et al., 2004*) et en champ visuel divisé (*i.e.* montrant de meilleures performances dans le champ visuel droit ; (*D'hondt & Leybaert, 2003*)). Concernant la moindre latéralisation de la tâche de catégorisation «comestible» par rapport à la tâche «vivant», nous suggérons que les items «comestibles» feraient intervenir une large composante émotionnelle de type «j'aime» ou «je n'aime pas» dans la prise de décision. En effet, dans une tâche de décision lexicale en champ visuel divisé sur des mots neutres et émotionnels, *Ortigue et al. (2004)* ont observé une dominance de l'hémisphère gauche quelque soit la valence émotionnelle des mots. Par ailleurs, ces auteurs ont démontré que cette latéralisation était moins importante pour les mots émotionnels que pour les mots neutres. Ils en ont ainsi conclu que le traitement des mots émotionnels engagerait non seulement l'hémisphère dominant gauche mais également, en parallèle, l'hémisphère droit. Cette étude pourrait alors expliquer nos résultats montrant une latéralisation réduite de la catégorisation «comestible» par rapport à la catégorisation «vivant». L'implication des deux hémisphères étant requise pour le traitement des mots émotionnels, cela conduirait à réduire la différence de performance entre les champs visuels.

3. Conclusion et perspectives

Nous avons tenté de pallier une limite importante de la technique en IRMf qui concernait la tâche utilisée. Même si un consensus paraît pour l'instant improbable, cette étude nous a permis de montrer que les tâches «classiquement» utilisées n'ont pas toutes le même «pouvoir» de dominance. Nous avons notamment pu remarquer que des variables psycholinguistiques telles que la valence d'imagerie ou émotionnelle des mots pouvait moduler la dominance hémisphérique du langage. A la suite de cette étude comportementale, nous avons réalisé une étude en IRMf en utilisant ces tâches. Deux résultats principaux se dégagent de cette étude IRMf : (i) concernant la représentation inter-hémisphérique, les deux tâches choisies induisent une latéralisation prégnante et (ii) concernant la représentation

² Tomographie par Émission de Positrons

intra-hémisphérique, ces tâches permettent de mettre en évidence les aspects sémantiques et phonologiques du langage classiquement décrits dans la littérature. Une perspective de ce travail est de mettre en place un véritable protocole constitué par un panel de tâches permettant une évaluation complète de la dominance hémisphérique du langage (représentation inter-hémisphérique) et de la représentation, au sein de l'hémisphère dominant, de zones dédiées à la phonologie et à la sémantique (intra-hémisphérique). Ce protocole devra être testé chez les sujets sains pour être ensuite appliqué chez les patients.

Références bibliographiques

- BACIU, Monica, LE BAS, Jean-François, SEGEARTH, Christoph & BENABID, Alim-Louis (2003). Presurgical fMRI evaluation of cerebral reorganization and motor deficit in patients with tumors and vascular malformations. *Eur J Radiol*, 46.2 : 139-46.
- BACIU, Monica, KAHANE, Philippe, MINOTTI, Lorella, CHARNALLET, Annick, DAVID, Danielle, LE BAS, Jean-François & SEGEARTH, Christoph (2001). Functional MRI assessment of the hemispheric predominance for language in epileptic patients using a simple rhyme detection task. *Epileptic Disorder*, 3.3 : 117-24.
- CAPITANI, Erminio, LAICONA, Mariani & MAHON, Bernard (2003). What are the facts of category-specific deficits ? A critical review of the clinical evidence. *Cognitive Neuropsychology*, 20.(3/6) : 213-61.
- CAPPA, Stefano, PERANI, Daniela, SCHNUR, Tatiana, TETTAMANTI, Marco & FAZIO, Ferruccio (1998). The effects of semantic category and knowledge type on lexical-semantic access: a PET study. *Neuroimage*, 8.4 : 350-9.
- D'HONDT, Murielle & LEYBAERT, Jacqueline (2003). Lateralization effects during semantic and rhyme judgement tasks in deaf and hearing subjects. *Brain Lang*, 87.2 : 227-40.
- EVIATAR, Zohar, HELLIGE, Joseph & ZAIDEL, Eran (1997). Individual differences in lateralization: effects of gender and handedness. *Neuropsychology*, 11.4 : 562-76.
- HELLIGE, Joseph (1996). Hemispheric asymmetry for visual information processing. *Acta Neurobiol Exp (Wars)*, 56.1 : 485-97.
- HERTZ-PANNIER, Lucie, GAILLARD, William Davis, MOTT, Stephen, CUENOD, Charles-André, BOOKHEIMER, Susan, WEINSTEIN, Steven, CONRY, Julianne, PAPERIO, Patricia, SCHIFF, Steven, LE BIHAN, Denis & THEODORE, William (1997). Noninvasive assessment of language dominance in children and adolescents with functional MRI: a preliminary study. *Neurology*, 48.4 : 1003-12.
- OLDFIELD, Richard (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9.1 : 97-113.
- ORTIGUE, Stéphanie, MICHEL, Christoph, MURRAY, Micah, MOHR, Christine, CARBONNEL, Serge & LANDIS, Théodore (2004). Electrical neuroimaging reveals early generator modulation to emotional words. *Neuroimage*, 21.4 : 1242-51.
- SEGHER, Mohamed, LAZEYRAS, François, PEGNA, Alan, ANNONI, Jean-Marie, ZIMINE, Ivan, MAYER, Eugène, MICHEL, Christophe & KHATEB, Asaid (2004). Variability of fMRI activation during a phonological and semantic language task in healthy subjects. *Hum Brain Mapp*, 23.3 : 140-55.
- SHAYWITZ, Bennett, SHAYWITZ, Sally, PUGH, Kenneth, CONSTABLE, Todd, SKUDLARSKI, Pawel, FULBRIGHT, Robert, BRONEN, Richard, FLETCHER, Jack, SHANKWEILER, Donald & KATZ, Leonard (1995). Sex differences in the functional organization of the brain for language. *Nature*, 373.6515 : 607-9.
- SPRINGER, Sally & DEUTSCH, Georg (2000). *Cerveau gauche, cerveau droit. A la lumière des neurosciences*. Bruxelles : De Boeck Université.
- TREMBLAY, Tania, MONETTA, Laura & JOANETTE, Yves (2004). Phonological processing of words in right- and left-handers. *Brain Cogn*, 55.3 : 427-32.

ÉTUDE DU LANGAGE SPONTANÉ D'ENFANTS SOURDS MOYENS ET LÉGERS : VERS UN ÉVITEMENT DE LA COMPLEXITÉ SYNTAXIQUE?

Hélène DELAGE

helene.delage@club-internet.fr

Université François Rabelais, Laboratoire Langage et handicap, Département de linguistique, UFR lettres et langues
37000 Tours

Abstract : Syntactic analysis was conducted on spontaneous language samples from children with mild-to-moderate hearing loss aged 6 to 11. Syntactic complexity measures show immaturity and differences in relative frequency of different types of subordinate clauses from that found in normal-hearing children. We propose an explanation related to mechanisms of avoiding complexity.

1. Introduction

1.1. Données de la littérature

Il existe peu d'études sur les effets linguistiques associés aux surdités moyennes et légères, pourtant plus fréquentes que les surdités plus sévères : en France, 420 000 enfants de 0 à 18 ans sont atteints par ce type de surdité, alors que les surdités sévères et profondes affectent 30 000 sujets dans cette même tranche d'âge (Gillot, 1998). Les difficultés scolaires et psychologiques (Davis, Elfenbein, Schum & Bentler, 1986) associées à la déficience auditive moyenne et légère sont très certainement tributaires des difficultés du développement du langage oral chez ces enfants. Les enfants sourds légers et moyens entrent dans le langage spontanément mais avec difficulté et avec du retard. Comme précisé dans Tuller & Jakubowicz (2004), ces enfants (à moins d'appartenir à une famille dont des membres ont des surdités plus sévères) n'apprennent pas une langue des signes et leur scolarité se déroule sans aide particulière (de type codage visuel). Leur seule entrée dans le langage est donc le langage oral, tel que perçu auditivement et visuellement *via* des informations obtenues au moyen de la lecture labiale. Le dépistage des surdités légères et moyennes reste très tardif (entre 4 et 5 ans selon les études) et donc, pour bon nombre d'enfants, la quasi-totalité de la période la plus cruciale pour l'acquisition du langage (la période dite « critique »), qui se situe avant l'âge de 6-7 ans (voir Bailey, Bruer, Symons & Lichtman, 2001 pour une revue de littérature), se déroule avec une perception de la parole lacunaire et déformée. Les quelques études consacrées à la description des performances langagières des enfants malentendants se fondent essentiellement sur des épreuves contraintes, c'est-à-dire en situation induite. Elles font apparaître un taux de troubles du langage oral bien supérieur à la norme, notamment dans le domaine de la morphosyntaxe (Norbury, Bishop & Briscoe, 2001, 2002 ; Tuller et Jakubowicz, 2004) et de la phonologie (Briscoe, Bishop & Norbury, 2001). Une très grande variabilité des performances langagières, non liée au degré de perte auditive, est mise en évidence dans ces études. L'étude du langage spontané permet d'appréhender le langage en situation plus naturelle (par rapport aux épreuves dites « contraintes »). Les productions linguistiques des sujets n'étant pas contraintes, ce contexte permet en effet d'appréhender de façon plus efficace d'éventuels moyens d'évitement. Les études que nous avons pu recenser sur le langage spontané dans le développement ordinaire à partir de 5 ans font apparaître une augmentation de la fréquence des structures syntaxiques complexes (Scott, 1988) et du taux de subordination chez des sujets âgés de 6 à 18 ans (Loban, 1976).

D'autres études, sur le développement atypique cette fois, ont relevé plusieurs traits caractéristiques (voir, par exemple, Reilly, Losh, Bellugi & Wulfeck, 2003) : des erreurs morphosyntaxiques plus fréquentes et un retard dans certaines mesures de complexité syntaxique, avec moins de subordonnées et une diversité moins riche (en ce qui concerne les différents types de constructions) que celle des enfants bénéficiant d'un développement ordinaire. Une seule étude s'est attachée à analyser le langage spontané d'enfants sourds moyens francophones (Guillemot, 2002). Des taux d'énoncés erronés importants avec des difficultés spécifiques à certaines catégories syntaxiques, comme les pronoms clitiques, ont été repérés mais aucune étude de la complexité syntaxique n'a été menée.

1.2. Problématique

L'objectif de notre étude est d'analyser la structure syntaxique du langage dans une population d'enfants sourds moyens et légers et d'identifier d'éventuels évitements de structures complexes. Pour ce faire, nous identifions tout d'abord les domaines langagiers déficitaires et les difficultés spécifiques puis nous affinons l'analyse en confrontant ces données, issues des épreuves contraintes, à celles issues de l'analyse du langage spontané. Si les performances en langage spontané se révèlent meilleures que celles obtenues dans les épreuves contraintes, ces dissociations pourraient être expliquées par la présence de moyens de compensation. Enfin, l'hétérogénéité des performances langagières étant une constante retrouvée dans la littérature, nous pourrions rapprocher cette variabilité de l'utilisation de stratégies différentes. Nous basant sur les résultats de Reilly *et al.* (2003) et Guillemot (2002), nous nous attendons à observer une complexité syntaxique moins grande dans notre population ainsi qu'un taux d'erreurs morphosyntaxiques important qui pourrait diminuer avec l'âge (un effet d'âge ayant été retrouvé dans la majorité des études précédemment citées sur la surdité moyenne). Les moyens de compensation, évoqués ci-avant, pourraient être caractérisés par un évitement de la subordination d'une part et par une prédominance des structures les moins complexes d'autre part. Nous entendons par structures moins complexes les subordonnées impliquant un calcul syntaxique moins coûteux, en l'occurrence les circonstancielles et les complétives non finies (hypothèses développées dans des communications récentes : Tuller, Delage & Monjauze, 2006 ; Hamann, Tuller, Delage, Henry & Monjauze, 2006).

2. Méthodologie

2.1. Participants

Notre étude vise à explorer la complexité syntaxique dans une population de 33 enfants sourds moyens et légers (notés 'ESML') âgés de 6;1 à 11;11 ans ($M = 9;2$ ans, $ET = 1;9$), originaires de la région Centre, tous porteurs d'une perte auditive neurosensorielle prélinguale bilatérale comprise entre 30 et 69 dB ($M = 43,7$ dB, $ET = 8,6$) et intégrés dans un cursus scolaire ordinaire. Vingt-trois enfants sont porteurs d'une surdité moyenne (comprise entre 41 et 69 dB), les 10 autres présentent une surdité légère (entre 30 et 40,9 dB). L'âge de dépistage est assez tardif ($M = 5;6$ ans, $ET = 1;5$). Les matrices progressives de Raven (Raven, Court & Raven, 1998), évaluant le niveau non verbal (capacités de raisonnement logique), ont été proposées à tous les enfants. Trente d'entre eux présentent un score supérieur au centile 10. Seuls 3 ont un score compris entre les centiles 5 et 10. Ces enfants étant intégrés en classe ordinaire, nous avons décidé de les maintenir dans notre population. Ont été exclus de l'étude tous les enfants présentant un contexte de bilinguisme familial ou un facteur associé à la surdité, tel qu'une déficience intellectuelle reconnue ou une surdité syndromique. Pour certaines mesures et lorsqu'il nous paraissait intéressant d'apprécier l'évolution des performances avec l'âge, nous avons scindé cette population en deux groupes d'âges : un groupe d'enfants âgés de 6;1 à 9;4 ans (notés 'ESML 6-9') et un groupe d'enfants

plus âgés entre 9;6 et 11;11 ans ('ESML 9-11'). Ces deux groupes ne se différencient pas par le degré de perte auditive.

En ce qui concerne l'analyse du langage spontané, « pivot » de cette étude, les performances des ESML sont comparées à celles de 3 groupes de 12 enfants normo-entendants âgés de 6, 8 et 11 ans et scolarisés en école primaire. Ces enfants témoins ont fait l'objet d'une sélection : exclusion des redoublants, des bilingues et des enfants suivis ou ayant été suivis pour des difficultés de langage (en orthophonie). Ils sont notés 'Témoins 6-11' ($M = 8;8$ ans, $ET = 2;2$) lorsqu'ils sont comparés dans leur ensemble aux ESML.

2.2. Matériels

Nous avons tout d'abord utilisé des épreuves dites contraintes (en situation induite) dans lesquelles sont ciblées des structures particulières. Un bilan complet du langage oral (BILO, Khomsi, sous presse) permet d'identifier les aires langagières les plus vulnérables. Ce bilan est informatisé et standardisé auprès d'une population de référence de 757 enfants et adolescents tout-venant. Les stimuli sont soit oraux (voix enregistrée), soit visuels (images présentées sur l'écran). Les aspects expressifs et réceptifs du langage oral sont évalués dans les domaines suivants :

- Compréhension orale (appariement phrase-image)
- Phonologie (répétition de mots)
- Vocabulaire (jugement lexical sur le versant réceptif, désignation d'images sur le versant expressif)
- Morphosyntaxe (jugement grammatical sur le versant réceptif, production d'énoncés sur le versant expressif)

Nous avons complété cette évaluation globale par un protocole expérimental : le PPPC (Protocole de Production des Pronoms Clitiques). Ce matériel permet d'induire la production d'éléments grammaticaux connus pour leur fragilité dans le contexte d'un développement atypique (Jakubowicz, Nash, Rigaut & Gerard, 1998), et notamment dans le contexte d'une surdité moyenne (Tuller & Jakubowicz, 2004 ; Delage & Tuller, sous presse) : les pronoms clitiques et tout particulièrement les clitiques accusatifs. Des images sont présentées sur un écran d'ordinateur, face à l'enfant, accompagnées de questions ciblées (présentées à l'oral par l'expérimentateur en même temps que l'image), comme en (1) où l'on attend la production d'un clitique nominatif (*il*) et d'un clitique accusatif à la 3^{ème} personne (*le*). Les réponses des enfants sont enregistrées puis transcrites et analysées à l'aide du système CHILDES (MacWhinney, 2000).

(1) Que fait le médecin avec le bébé ?

Réponse attendue : il le pèse.

Les données issues de ces deux épreuves contraintes ont pour but d'être comparées aux résultats obtenus lors de l'analyse syntaxique du langage spontané. Des échantillons de 50 à 70 énoncés (en moyenne 58 énoncés) de plus d'un mot ont été enregistrés et transcrits pour chaque enfant à l'aide du système de transcription CHILDES. Le recueil consiste en une conversation libre entre l'expérimentateur et l'enfant sur des thèmes variés (loisirs, école, récits de films ou livres...). Le calcul de la longueur moyenne d'énoncés (=nombre moyen de mots par énoncé) et du taux d'erreurs (=N erreurs morphosyntaxiques/N total d'énoncés) permettent une première analyse syntaxique. Notre analyse du langage spontané cible plus particulièrement la complexité syntaxique des énoncés *via* différentes mesures :

- Taux de subordination : nombre de subordonnées / nombre d'énoncés verbaux
- Taux d'enchâssement profond : nombre de subordonnées enchâssées dans une autre subordonnée / nombre de propositions, comme en (2)

- Diversité de subordination : répartition des différents types de propositions subordonnées parmi les relatives (3), circonstancielles (4), complétives finies (5) et non finies (6) c'est-à-dire sans flexion verbale.
- (2) Je vais t'en montrer un [parce que je crois [que je l'ai l'image]]. (10;3 ans)
- (3) Y a une dame [qu(i) arrive]. (11;5 ans)
- (4) Si parfois ils jouent mais [quand ils jouent à ça] je leur casse les pieds. (6;7 ans)
- (5) Elle aime bien [que je tape sur ses doudous]. (11;6 ans)
- (6) Faut [lui donner une fessée] hein ! (7;4 ans)

3. Évaluations en situation contrainte

Afin de constater la présence de troubles langagiers chez les ESML, il est nécessaire de fixer un seuil à partir duquel ces troubles peuvent être identifiés. Dans la littérature, il n'existe pas de consensus pour caractériser un trouble langagier, les seuils se situant entre -1 et -2 Ecart-Type suivant les études. Nous avons choisi d'adopter le seuil de -1,65 ET, seuil utilisé en clinique et correspondant au 5^{ème} percentile dans une population normale, pour identifier des troubles langagiers patents. Les domaines langagiers déficitaires, mis en évidence par les différentes épreuves du BILO (présentées précédemment), sont représentés sur la figure 1 pour les ESML 6-9 ans et 9-11 ans.

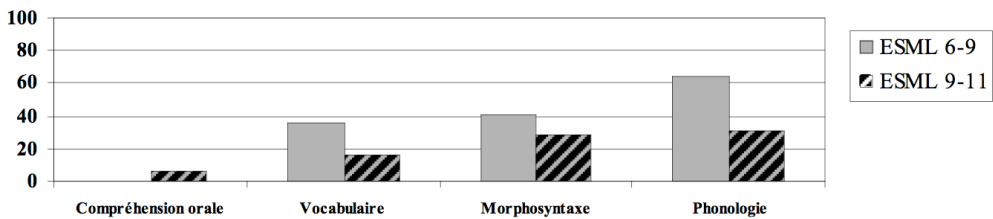


Figure 1 - Domaines langagiers déficitaires : % de scores < -1,65 ET

Ces résultats mettent en évidence des troubles langagiers prédominant en morphosyntaxe et en phonologie (déficits les plus cités dans la littérature). Dans une visée développementale, et bien que nous ne disposions que de données transversales, il semble que les troubles du vocabulaire et de la phonologie diminuent davantage avec l'âge que les troubles de la morphosyntaxe qui seraient donc plus persistants.

Pour ce qui est du protocole ciblé sur la production des pronoms clitiques, nous avons mis en évidence, comme attendu, des difficultés dans la production du clitique accusatif (à la 3^{ème} personne). Ainsi, pour cette variable syntaxique, considérée dans la littérature comme un marqueur de dysfonctionnement langagier en français (voir par exemple Jakubowicz *et al.*, 1998 ; Paradis & Crago, 2001), les ESML 6-9 ont un taux de production (50%) significativement inférieur à celui des témoins de 6 ans (74,2%) ($U = 57, p < .05$). De même, les ESML 9-11 ont un taux de production (75%) inférieur à celui des témoins de 11 ans (98%) ($U = 39, p < .01$), mais le fait le plus intéressant réside dans la similarité des réponses produites par les ESML 9-11 et les témoins âgés de 6 ans. Le tableau 1 présente les différents types de réponses obtenues lors de la production induite du pronom clitique accusatif pour ces deux populations. Au niveau syntaxique, les réponses contenant un syntagme nominal (exemple : *il pèse le bébé*) sont correctes mais discursivement inadéquates, le patient étant déjà exprimé dans la question (*Que fait le médecin avec le bébé ?*). Les omissions (de clitiques accusatifs) constituent, du moins dans le contexte de ce protocole, des réponses agrammaticales (exemple : *il pèse*). Enfin, les réponses « non cible » ne correspondent pas à la cible syntaxique attendue (exemple de réponse : *il regarde combien il pèse*).

	% de production	Production de syntagmes nominaux	Omissions	Réponses non cibles
ESML 9-11	75%	11,5%	7,3%	6,2%
Témoins 6 ans	74,2%	12,5%	8,6%	4,7%

Tableau 1 - Productions induites du pronom accusatif 3ème personne

A la lumière de ces résultats, les réponses des ESML 9-11 et celles des témoins de 6 ans suivent un pattern identique. Ce rapprochement évoque davantage un retard dans le développement langagier des ESML qu'un dysfonctionnement spécifique à la surdité.

4. Analyse syntaxique du langage spontané

4.1. Mesures de complexité

Le tableau 2 expose les différents résultats issus de l'analyse du langage spontané (avec la longueur moyenne d'énoncés et les taux d'erreurs) et plus particulièrement de la complexité syntaxique avec les taux de subordination et d'enchâssement profond.

	Longueur moyenne d'énoncés	Taux d'erreurs	Taux de subordination	Taux d'enchâssement profond
ESML 6-9	5,7	14,5%	19,3%	1,2%
ESML 9-11	6,7	9,2%	30,1%	3,2%
Témoins 6 ans	7,1	2,5%	32%	2,9%
Témoins 8 ans	7,7	1,6%	35,8%	3,7%
Témoins 11 ans	8,1	0,8%	39,7%	3,3%

Tableau 2 - Mesures syntaxiques issues de l'analyse du langage spontané (moyennes par groupe d'âge)

Pour la quasi-totalité des mesures, les résultats des ESML sont inférieurs à ceux des témoins. La longueur moyenne d'énoncés augmente significativement avec l'âge chez les ESML (entre ESML 6-9 et 9-11 : $U = 68$, $p < .05$) mais reste statistiquement inférieure à celle des témoins (ESML 6-9 < Témoins 6 ans : $U = 33$, $p < .01$; ESML 9-11 < Témoins de 8 ans : $U = 45$, $p < .05$). Les taux d'erreurs des ESML sont, comme on s'y attendait, particulièrement élevés et significativement plus importants que ceux des témoins ($U = 85$, $p < .001$). En revanche, on ne trouve pas d'effet d'âge significatif sur cette mesure chez les ESML, ce qui s'explique par la très grande variabilité retrouvée dans cette population. Ainsi, chez les ESML 6-9, les taux d'erreurs varient entre 3 et 56% ; l'étendue est également importante chez les ESML 9-11 (entre 2 et 27%). Enfin, il faut noter que ces erreurs morphosyntaxiques touchent les pronoms à hauteur de 38% (omissions, erreurs de genre, de nombre ou de référence). Les deux mesures de complexité syntaxique (taux de subordination, taux d'enchâssement profond) donnent des résultats sensiblement différents : des performances faibles pour les ESML 6-9 (avec des taux statistiquement inférieurs à ceux des témoins de 6 ans) et des performances proches de celles des témoins pour les ESML 9-11.

4.2. Vers un évitement de la complexité syntaxique ?

Nous avons évoqué dans la section précédente le fait que les taux de subordination et d'enchâssement profond des ESML 9-11 se rapprochaient de ceux des témoins. Les ESML 9-11 ont néanmoins un taux d'erreurs significativement plus important et des troubles de la morphosyntaxe avérés en situation contrainte pour 28% d'entre eux (seuil de -1,65 ET) ou, si l'on choisit un seuil moins restrictif, pour 41% des sujets (seuil de -1 ET). Les performances langagières des ESML sont également caractérisées par un retard dans la production du clitique accusatif (3^{ème} personne), marqueur de dysfonctionnement langagier. Nous souhaitons mettre en rapport les données issues des évaluations contraintes avec celles issues de l'évaluation « naturelle » et nous ne trouvons pas de corrélation entre les mesures de langage spontané et les épreuves standardisées de morphosyntaxe.

Ces observations nous amènent à nous interroger sur la présence de moyens de compensation dans le langage spontané. Les ESML parviennent à produire des phrases complexes (même les plus jeunes en produisent, certes moins souvent que les témoins du même âge) ; nous pouvons imaginer qu'ils utilisent, parmi les différents types de subordonnées, les structures les moins complexes. Nous avons donc analysé la diversité de subordination, c'est-à-dire la répartition des différents types de subordonnées. La figure 2 illustre cette diversité chez les témoins 6-11 et les ESML 6-11.



Figure 2 - Répartition des types de subordonnées

Cette figure met en lumière des patterns différenciant les deux populations, les ESML produisant davantage de complétives non finies que les témoins ($U = 404, p < .05$). Il semble donc qu'il y ait évitement de l'accord sujet et temporel chez les ESML qui seraient particulièrement sensibles aux calculs syntaxiques impliqués dans la flexion verbale.

Après avoir analysé la diversité syntaxique entre les subordonnées, nous pouvons nous interroger sur la répartition des différents types de structures à l'intérieur même des subordonnées. Ainsi, parmi les relatives, les relatives objets (dont le pronom est en position syntaxique d'objet) sont plus difficiles à traiter (Gibson, 1998) et plus complexes que les relatives sujets qui impliquent un déplacement de moins grande distance (Tuller *et al.*, 2006 ; Hamann *et al.*, 2006). Les enfants de l'étude sont sensibles à cette différence de complexité car, chez les ESML comme chez les témoins, la production de relatives sujets est significativement plus importante que celle des relatives objets ($p < .001$). Comparer les deux populations est alors particulièrement intéressant : le taux de relatives objets produites par les ESML est significativement inférieur à celui des témoins 6-11 ($U = 404, p < .05$) ce qui n'est pas le cas pour les relatives sujets ($p = 0.9$) ; on retrouve donc un évitement de la complexité syntaxique chez les enfants malentendants. L'étude des productions individuelles renforce cette différence : seulement 36% des ESML produisent au moins une relative objet contre 69% des témoins. Une grande majorité des ESML semble donc éviter la production des structures syntaxiques les plus complexes.

4.3. La production d'erreurs est-elle liée à la complexité syntaxique ?

Les observations précédentes vont dans le sens d'un évitement des structures complexes chez les ESML. Nous avons également souligné que les ESML, quel que soit leur âge, se distinguaient des témoins par un taux d'erreurs très élevé. Il est tout à fait logique d'imaginer que la production d'énoncés complexes engendre, chez ces enfants particulièrement sensibles à la complexité, davantage d'erreurs (hypothèse déjà confirmée par une analyse similaire chez des adolescents dysphasiques, Henry, 2006). Nous avons donc calculé le pourcentage d'erreurs portant sur les énoncés complexes en le comparant à celui portant sur les énoncés simples. Les ESML, mais aussi les témoins, sont effectivement sensibles à la complexité et produisent davantage d'erreurs sur les énoncés complexes, et ceci en dépit du fait que les témoins ont des taux d'erreurs très faibles. En revanche, cette différence est statistiquement plus forte chez les ESML ($p < .001$) que chez les témoins ($p < .05$). Afin d'approfondir cette analyse, nous avons calculé la répartition des erreurs au sein des

différents types de subordonnées. Autrement dit, les ESML font-ils davantage d'erreurs sur les subordonnées les plus complexes ? La figure 3 présente cette répartition avec les pourcentages d'erreurs portant sur les relatives, les complétives finies, les circonstancielles et les complétives non finies qui, nous le rappelons, ne nécessitent pas d'accord.

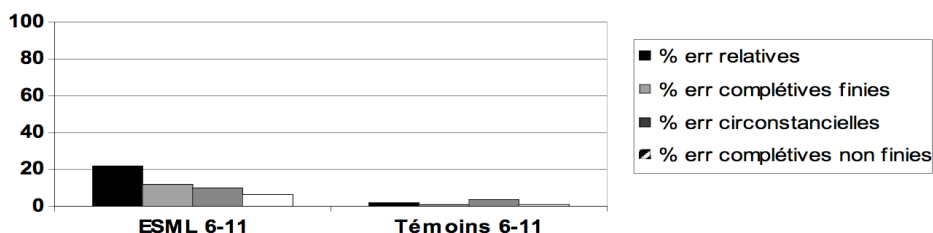


Figure 3 - Répartition (%) des erreurs suivant le type de subordonnée

Cette figure met là encore en avant la grande sensibilité des ESML à la complexité syntaxique. En effet, les erreurs portant sur les relatives, considérées comme particulièrement complexes, sont significativement plus fréquentes que les erreurs portant sur les trois autres types de subordonnées ($T = 45$, $p < .05$). D'autre part, les ESML qui, on l'a souligné, produisent davantage de complétives non finies, font peu d'erreurs sur ce type de subordonnée. Les moyens de compensations utilisés paraissent donc « judicieux » : les ESML produisent en majorité les structures les moins complexes et donc les moins sujettes à erreur.

5. Conclusion

Nous avons mis en évidence chez les ESML un schéma de développement typique mais retardé, avec une amélioration des performances avec l'âge dans les épreuves standardisées ainsi qu'une augmentation de la complexité syntaxique. Nous avons également démontré la présence de moyens d'évitement de la complexité syntaxique, caractérisés par une prédominance des structures les moins complexes (ne nécessitant pas de flexion verbale). D'autre part, les performances en langage spontané des ESML se caractérisent par des taux d'erreurs élevés, et notamment pour les structures complexes comme les relatives. Ces résultats nous poussent à croire qu'on trouvera, chez des adolescents, davantage de moyens de compensation, phénomène qui serait lié à une diminution des erreurs (Delage, étude longitudinale en cours) ; diminuer la complexité équivaldrait alors à minimiser les sources d'erreurs.

Références

- BAILEY, Donald B., BRUER, John T., SYMONS, Frank J., & LICHTMAN, Jeff W. (Eds) (2001). *Critical thinking about critical periods. A series from the national center for early development and learning*. Baltimore: Brookes Publishing.
- BRISCOE, Josie, BISHOP, Dorothy & NORBURY, Courtenay-Frazier (2001). Phonological processing, language, and literacy: a comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment, *Journal of Child Psychology and Psychiatry*, 42: 329-340.
- DAVIS, Julia-M., ELFENBEIN, Jill, SCHUM, Robert & BENTLER, Ruth-A. (1986). Effects of mild and moderate hearing impairment on language, educational, and psychosocial behavior of children, *Journal of Speech and Hearing Disorders*, 51: 56-62.
- DELAGE, Hélène & TULLER, Laurice (sous presse). Language development and mild-to-moderate hearing loss: Does language normalize with age?, *Journal of Speech, Language, and Hearing Research*.

- GIBSON, Edward (1998). Linguistic complexity: locality of syntactic dependencies, *Cognition*, 68: 1-76.
- GILLOT, Dominique (1998). *Le droit des sourds : 115 propositions*. Rapport parlementaire du 30 juin 1998.
- GUILLEMOT, G r aldine (2002). *Etude des cat gories morphosyntaxiques dans le langage spontan  de douze enfants sourds moyens*. M moire de capacit  en orthophonie et de ma trise de sciences du langage, Universit  Fran ois Rabelais, Tours.
- HAMANN, Corn lia, TULLER, Laurice, DELAGE, H l ne, HENRY, C lia & MONJAUZE, C cile (2006). (Un)successful subordination in French-speaking children and adolescents with SLI. Communication orale. *The Romance Turn II*, Utrecht, September 7-9.
- HENRY, C lia (2006). *Compensation chez les adolescents dysphasiques : Analyse du langage spontan *. M moire de D.E.A., Universit  Fran ois Rabelais, Tours.
- JAKUBOWICZ, C lia, NASH, L a, RIGAUT, Catherine & GERARD, Christophe-Lo c (1998). Determiners and Clitic Pronouns in French-Speaking Children with SLI, *Language Acquisition*, 7: 113-160.
- KHOMSI, Abdelhamid (sous presse). *Bilan Informatis  du Langage Oral*. Paris: ECPA.
- LOBAN, Walter (1976). *Language development: kindergarten through grade twelve*. Urbana, IL: National Council of Teachers of English.
- MACWHINNEY, Brian (2000). *The CHILDES project: Tools for analyzing talk*. 3e  dition. Mahwah, NJ : Lawrence Erlbaum Associates.
- NORBURY, Courtenay-Frazier, BISHOP Dorothy & BRISCOE Josie (2001). Production of english finite verb morphology: A comparison of SLI and mild-moderate hearing impairment, *Journal of Speech, Language, and Hearing Research*, 44: 165-179.
- NORBURY, Courtenay-Frazier, BISHOP, Dorothy & BRISCOE, Josie (2002). Does impaired grammatical comprehension provide evidence for an innate grammar module?, *Applied Psycholinguistics*, 23: 247-268.
- PARADIS, Johanne & CRAGO, Martha (2001). The morphosyntax of Specific Language Impairment in French: An extended optional default account, *Language Acquisition*, 9: 269-300.
- RAVEN, John-Carlyle, COURT, J.H. & RAVEN, J. (1998). *Progressive matrices Couleur*. Paris : Editions et applications Psychologiques.
- REILLY, Judy, LOSH, Molly, BELLUGI, Ursula & WULFECK, Beverly (2003). "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome, *Brain and Language*, 88: 229-247.
- SCOTT, Ch ryl-M. (1988). Spoken and written syntax, in *Later Language development: Ages 9 through 19*, Nippold, Marilyn (Ed). San Diego, CA: College Hill Press, 44-95.
- TULLER, Laurice, DELAGE, H l ne & MONJAUZE, C cile (2006). Avoiding complexity in atypical development of French, *Latsis Colloquium of the University of Geneva*, January 26-28.
- TULLER, Laurice & JAKUBOWICZ, C lia (2004). D veloppement de la morphosyntaxe du fran ais chez des enfants sourds moyens, *Le Langage et l'Homme : Logop die, Psychologie, Audiologie*, 14, 2: 191-207.

LES TROUBLES PRAGMATIQUES ET ÉMOTIONNELS CHEZ DES ENFANTS ET ADOLESCENTS CÉRÉBROLÉSÉS : L'EXEMPLE DE LA SITUATION DE PRODUCTION NARRATIVE

Anaïg DELANOË

anaig.delanoie@uhb.fr

Université Rennes II – Centre de Recherche en Psychologie, Cognition et Communication

Place du Recteur Henri Le Moal – CS 24307

35043 Rennes Cedex

France

Abstract : The main goal of this study is to analyse social uses of language in situations of communication. The performances of six children and adolescents with cerebral damage were compared to those of six age-matched controls in a narrative production task. The results show that these patients have pragmatic and emotional difficulties in carrying out the task.

1. Introduction

Des recherches consacrées aux compétences pragmatiques ont été réalisées chez des adultes cérébrólésés. Ces recherches récentes ont montré le rôle important de l'hémisphère droit (Champagne, Virbel, Nespoulous et Joannette, 2003 ; Joannette, Goulet et Hannequin, 1990 ; Stemmer et Cohen, 2002) et des lobes frontaux (Bernicot et Dardier, 2001 ; Dardier, 2001 ; McDonald et Pearce, 1998 ; Peter-Favre, 1995, 1999) dans le traitement du langage en contexte. Néanmoins, même si des troubles pragmatiques sont désormais décrits chez des patients cérébrólésés adultes, il existe encore peu d'études consacrées aux enfants et aux adolescents cérébrólésés (Dardier, 2004). Or, les troubles pragmatiques constituent d'importants obstacles en terme d'adaptation et d'intégration sociale des enfants et des adolescents souffrant de lésions frontales (Dardier, 2004). La compréhension de ces troubles est donc nécessaire tant sur le plan théorique que clinique, afin de mieux appréhender les troubles des patients et améliorer leur prise en charge (Peter-Favre, 2002). Enfin, il existe peu de données et d'outils à disposition des chercheurs et des cliniciens dans le domaine de l'évaluation des troubles pragmatiques. Il apparaît actuellement nécessaire de décrire ces troubles et les dissociations éventuelles rencontrées dans les profils pragmatiques, afin d'envisager, entre autres, la construction de tels outils. D'où l'intérêt de cette recherche préliminaire qui aborde l'analyse des compétences pragmatiques auprès d'une population d'enfants et d'adolescents cérébrólésés.

L'objectif principal de notre travail est d'étudier le fonctionnement des usages sociaux du langage en situation de communication. Dans cette recherche, nous analysons les performances d'enfants et d'adolescents cérébrólésés et de sujets tout-venant dans une épreuve de production narrative. La tâche de production narrative est intéressante pour plusieurs raisons. D'une part, le récit est un genre culturellement déterminé (Berman et Slobin, 1994 ; Kern, 1997, 2002). D'autre part, cette tâche permet l'analyse d'une grande variété de caractéristiques linguistiques, formelles (phonologie, syntaxe, etc.) et pragmatiques (Reilly et Wulfeck, 2004). Enfin, le récit est également utilisé pour étudier d'autres domaines cognitifs, comme l'expression, la compréhension et l'attribution des émotions (Bellugi, Reilly, Krieter, Doyle et Jones, 2003 ; Bernicot, Lacroix et Reilly, 2003). Selon nous, compte tenu de l'ensemble de ses spécificités, la situation de narration est une tâche complexe qui peut donc constituer une « épreuve frontale » sollicitant les fonctions

exécutives (par exemple, le traitement séquentiel, l'élaboration d'un plan, la flexibilité mentale, l'inhibition de réponses non pertinentes, le maintien de la stratégie et la vérification du résultat final) ; cette épreuve peut permettre de mettre en évidence des difficultés variées dans l'utilisation du langage chez les patients cérébrolésés.

La littérature relative aux troubles pragmatiques observés chez les patients cérébrolésés souligne l'existence de dissociations dans les performances des patients, en particulier des patients lésés frontaux (Dardier, Delaye et Laurent-Vannier, 2003). Grâce à cette étude, nous verrons si des dissociations existent dans les résultats des patients à l'épreuve de narration.

Les recherches antérieures réalisées auprès d'adultes et d'enfants cérébrolésés soulignent leurs difficultés à produire des récits. Dans ce domaine, il n'existe pas de travaux consacrés spécifiquement aux troubles du récit chez les enfants et les adolescents. Néanmoins, des études consacrées à l'analyse de leurs capacités conversationnelles évoquent leurs difficultés dans cette activité (en ce qui concerne la gestion des thèmes ou la relation à l'interlocuteur) et soulignent le rôle majeur des fonctions exécutives dans les situations de communication qui nécessitent l'application de règles d'usages du langage spécifiques (Bernicot, Dardier, Fayada, Pillon et Dubois, 2001 ; Dardier, 2001, 2004 ; Dardier et Bernicot, 2000 ; Peter-Favre, 1995, 1999). Compte tenu de l'ensemble de ces recherches, nous faisons l'hypothèse qu'à l'inverse des sujets contrôles, les patients cérébrolésés vont montrer des difficultés dans une épreuve de narration qui sollicite des compétences linguistiques et pragmatiques variées (Reilly et Wulfeck, 2004). Ces difficultés s'observeront notamment sur le plan de la structuration du récit, car les patients présenteront une réduction du discours et des troubles de la cohésion. Au niveau pragmatique, ils auront plus de difficultés que les sujets contrôles à respecter le schéma canonique et la continuité thématique de l'histoire. Enfin, compte tenu des troubles du comportement fréquemment décrits chez les patients cérébrolésés, en particulier chez les patients lésés frontaux, nous faisons l'hypothèse qu'un trouble dans la gestion des comportements émotionnels sera mis en évidence chez ces patients par l'analyse des évaluations sociales (intensificateurs, discours direct, questions et exclamations dirigées vers l'expérimentateur).

2. Méthodologie

2.1. Sujets

La population est composée de deux groupes d'enfants et d'adolescents, de langue maternelle française : un groupe de patients souffrant de lésions cérébrales acquises (unilatérales ou bilatérales) et un groupe de sujets contrôles. Tous étaient volontaires pour participer à cette recherche. Le groupe expérimental est composé de six enfants et adolescents, suivis au Centre de Rééducation Fonctionnelle de Kerpape, dans le Morbihan et à l'hôpital Saint-Maurice, à Paris. Les patients sont âgés de 8 ans 6 mois à 20 ans 2 mois (quatre garçons et deux filles). Ils ne présentent pas de déficience intellectuelle majeure (le QI total, mesuré à l'aide du WISC-III, est supérieur à 70), ni de troubles aphasiques ou d'autres troubles susceptibles de nuire à la passation. Les caractéristiques complètes des patients se trouvent dans le tableau 1. Le groupe contrôle est composé de six enfants et adolescents, scolarisés en région Bretagne (non redoublants), appariés aux patients au niveau du sexe et de l'âge (plus ou moins trois mois). Les sujets contrôles sont âgés de 8 ans 8 mois à 20 ans 3 mois (quatre garçons et deux filles).

Patient	Sexe	Age	Délai entre la lésion et l'étude	Quotient Intellectuel Total	Localisation lésion	Etiologie lésion
F1	Masculin	8 ans 6 mois	4 ans 8 mois	71	Frontal – Temporal – Occipital	Traumatisme crânien
F2	Masculin	9 ans 10 mois	9 mois	97	Frontal - Occipital	Traumatisme crânien
F3	Masculin	12 ans 6 mois	8 ans 2 mois	85	Frontal	Tumeur
F4	Masculin	16 ans 10 mois	7 mois	77	Frontal – Temporal	Traumatisme crânien
F5	Féminin	18 ans 2 mois	1 an 8 mois	74	Frontal	Traumatisme crânien
F6	Féminin	20 ans 2 mois	8 mois	113	Frontal - Temporal	Traumatisme crânien

Tableau 1 - Caractéristiques des sujets du groupe expérimental

2.2. Matériel et procédure

Le matériel est constitué d'un livre intitulé « Frog, where are you ? » (Mayer, 1969), composé de vingt-quatre images en noir et blanc sans texte. Ce livre « raconte » l'histoire d'un petit garçon qui, ayant perdu sa grenouille, part à sa recherche et la retrouve après une série d'aventures. La figure 1 présente trois images constituant le livre.

Les enfants regardent une première fois les images puis ils doivent raconter l'histoire en les regardant à nouveau.



Figure 1 - Trois des vingt-quatre images faisant partie de l'épreuve de production narrative

2.3. Indices utilisés

Pour l'ensemble des indices, le codage est effectué en aveugle par deux juges et l'accord entre les deux juges est supérieur à 80%. Les indices utilisés dans cette étude peuvent être regroupés en deux catégories : d'une part, les indices pragmatiques et d'autre part, les évaluations sociales. Dans le cadre de l'étude des compétences pragmatiques, nous évaluons la façon dont les sujets développent et structurent leur récit en analysant le nombre de propositions, les pourcentages de pronoms ambigus et de phrases complexes. Nous évaluons également la façon dont les sujets respectent le schéma d'une histoire culturellement transmis (analyse du score du schéma canonique sur une échelle en 14 points, d'après Bernicot *et al.*, 2003) et dont ils développent et maintiennent le thème du récit (analyse du score de continuité thématique sur une échelle en 5 points, d'après Bernicot *et al.*, 2003 ; Reilly, Losh, Bellugi et Wulfeck, 2004). Dans le cadre de l'étude des évaluations sociales, nous évaluons la position plus ou moins subjective du locuteur par rapport à l'auditeur et à l'histoire (Bernicot *et al.*, 2003 ; Losh, Bellugi, Reilly et Anderson, 2000 ; Reilly *et al.*, 2004), en relevant les marques de discours direct (par exemple : « Le petit garçon dit : "Oh ! La belle grenouille !" »), les questions et exclamations produites vers l'expérimentateur (par exemple : « Regarde ! Il a des bois le cerf ! »), ainsi que les intensificateurs (par exemple : « Il est trop en colère le petit garçon. »).

3. Résultats

Les différents résultats ont été traités à l'aide du test non-paramétrique U de Mann-Whitney. Les différences ont été considérées comme étant significatives à partir du seuil de .05.

3.1. Aspects pragmatiques

Les figures 2 à 4 présentent les résultats des sujets des deux groupes en ce qui concerne le nombre de propositions, le pourcentage de pronoms ambigus et de phrases complexes (par rapport au nombre total de propositions).

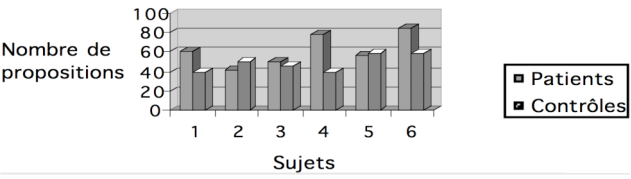
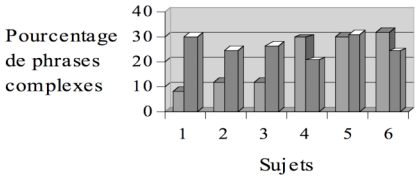
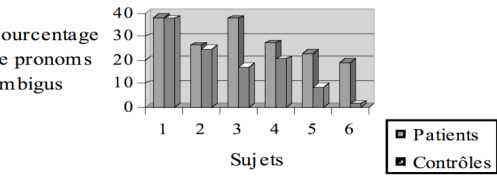


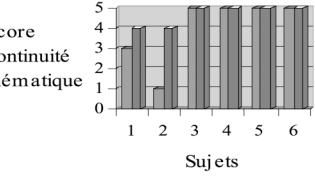
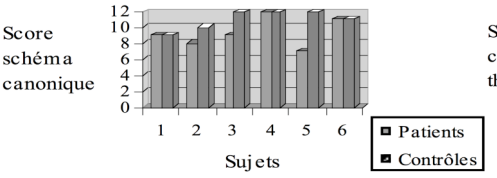
Figure 2 - Nombre de propositions constituant les récits des sujets des deux groupes



Figures 3 et 4 - Pourcentages de pronoms ambigus et de phrases complexes constituant les récits des sujets des deux groupes

L'effet du groupe n'est pas significatif pour le nombre de propositions et le pourcentage de phrases complexes. En revanche, le pourcentage de pronoms ambigus produit par le groupe de patients est significativement plus élevé que celui du groupe contrôle ($U(6;6)=28$; $p<.05$).

Les figures 5 et 6 indiquent les scores pour les sujets des deux groupes en ce qui concerne le schéma canonique et le respect de la continuité thématique.

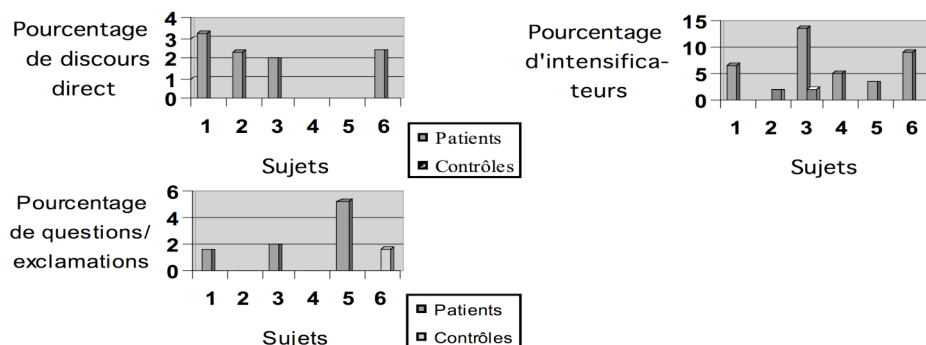


Figures 5 et 6 - Scores de respect du schéma canonique et de la continuité thématique des récits des sujets des deux groupes

Pour l'indice « schéma canonique », l'effet du groupe est significatif ($U(6;6)=8$; $p<.05$). Les sujets contrôles respectent mieux le schéma de l'histoire que les patients cérébrolésés. En revanche, pour l'indice « continuité thématique », l'effet de groupe n'est pas significatif.

3.2. Aspects émotionnels

Les figures 7, 8 et 9 indiquent les résultats des sujets des deux groupes en ce qui concerne le pourcentage de production d'évaluations sociales.



Figures 7, 8 et 9 - Pourcentages d'évaluations sociales constituant les récits des deux groupes

Pour le pourcentage de discours direct, l'effet du groupe est significatif ($U(6;6)=30$; $p<.01$). De même, en ce qui concerne la production d'intensificateurs, l'effet du groupe est significatif ($U(6;6)=32$; $p<.01$). Les patients cérébrolésés produisent davantage de discours direct et d'intensificateurs que les sujets contrôles. En revanche, l'effet du groupe n'est pas significatif pour le pourcentage de questions et exclamations envers l'expérimentateur.

4. Discussion

Rappelons tout d'abord que l'objectif de ce travail était d'effectuer une recherche préliminaire concernant les troubles pragmatiques et émotionnels chez des enfants et adolescents cérébrolésés. Pour ce faire, nous avons mis en place un protocole expérimental composé d'une tâche de production de récit. Compte tenu de la taille restreinte de notre population expérimentale et de son hétérogénéité en terme d'âge, nous ne pouvons conclure définitivement ici sur les capacités pragmatiques des patients. Néanmoins, ce premier travail exploratoire souligne l'existence de difficultés subtiles chez cette population.

Ainsi, la première question posée était de savoir si dans une épreuve de narration, les patients cérébrolésés non aphasiques avaient des récits différents des sujets contrôles. Plus précisément, deux niveaux dans la narration ont été étudiés : un niveau pragmatique (mesure du nombre de propositions, des pourcentages de phrases complexes et de pronoms ambigus, mesure du respect du schéma canonique et de la continuité thématique) et un niveau émotionnel (mesure de la régulation des émotions dans le récit).

Concernant les aspects pragmatiques du récit, les résultats obtenus confirment partiellement notre hypothèse de départ. En effet, l'analyse du nombre de propositions ne montre pas de différences entre les récits des enfants et adolescents cérébrolésés et ceux des sujets contrôles. Ces résultats rejoignent ceux obtenus auprès d'enfants et d'adolescents cérébrolésés dans une situation de conversation (Dardier, Reilly, Bates, Delaye et Laurent-Vannier, 2004), mais vont à l'encontre d'autres travaux qui évoquent une réduction du discours chez des enfants et adolescents lésés frontaux en situation de récit (Chapman, Levin, Wanek, Weyrauch et Kufera, 1998 ; Dennis, Jacennik et Barnes, 1994 ; Reilly, Bates et Marchman, 1998).

L'analyse de la cohésion des récits montre que les productions des enfants et des adolescents cérébrolésés sont moins organisées que celles des sujets contrôles, car ils utilisent davantage de pronoms ambigus et explicitent donc moins les liens existant entre les unités linguistiques et leurs référents. Ces résultats confirment ceux obtenus par Chapman *et al.* (1992), Dardier *et al.* (2004) et Dennis *et al.* (1994) auprès d'enfants et d'adolescents souffrant de lésions frontales. Dans la mesure où les unités linguistiques ne peuvent être considérées indépendamment de leur valeur fonctionnelle au sein du récit, les troubles de la cohésion narrative des enfants et adolescents cérébrolésés peuvent être mis en lien avec leurs déficits

en terme de capacités d'adaptation à une situation de communication. Dans cette épreuve, c'est bien l'organisation et la structuration du récit qui apparaît déficitaire chez les patients.

En outre, les récits des patients cérébrolésés montrent aussi des spécificités par rapport à ceux des sujets contrôles. Ainsi, nos résultats mettent en évidence des difficultés chez ces patients à respecter une règle fondamentale lors de la production des récits : le respect du schéma canonique. Sur ce point, nos données rejoignent les résultats déjà obtenus auprès d'adultes cérébrolésés dans des situations de récits par Biddle, McCabe et Bliss (1996), Coelho, Liles et Duffy (1995), Glosser et Deser (1990), Liles, Coelho, Duffy et Zalagens (1989) ou par Mentis et Prutting (1991) et Peter-Favre (1995 ; 1999 ; 2002) dans des situations de conversations. En revanche, nos résultats ne sont pas en accord avec ceux obtenus par ces auteurs en ce qui concerne le respect de la continuité thématique. Dans le cadre de notre étude, les performances des patients ne sont pas inférieures à celles des contrôles. Selon nous, il est possible que cette absence de différence soit liée à l'effet plafond observé chez les sujets contrôles les plus âgés dans le cadre de l'indice de continuité thématique. Sur ce point, une première critique peut être effectuée : en effet, une évaluation plus complexe de cet indice, intégrant davantage d'éléments à restituer permettrait sans doute dans l'avenir de mieux différencier les récits des patients de ceux des sujets contrôles.

Selon nous, il est possible de faire un premier lien entre ces troubles pragmatiques dans la gestion de la structure du récit et des troubles des fonctions exécutives, dans la mesure où tous nos patients présentent des lésions cérébrales incluant les régions frontales. En effet, la production d'un récit peut être considérée comme une tâche complexe sollicitant les fonctions exécutives, dans la mesure où elle nécessite la mise en œuvre de différentes étapes pour sa réalisation. Nous rejoignons le point de vue de Mar (2004) pour lequel les lobes frontaux et par conséquent les fonctions exécutives sous-jacentes, sont sollicités dans notre épreuve de narration. Parmi ces fonctions exécutives, nous pouvons notamment citer le traitement séquentiel des informations pertinentes dans le récit, l'élaboration et l'organisation d'une trame narrative, l'inhibition des informations non pertinentes. Sur ce point, des recherches complémentaires incluant une évaluation neuropsychologique complète des patients sont désormais nécessaires afin d'affiner l'analyse des traitements inférentiels mis en œuvre dans ces situations de communication et aussi pour évaluer l'influence dans ce domaine d'autres troubles neuropsychologiques.

En ce qui concerne les aspects émotionnels des narrations, notre analyse a porté sur la régulation des émotions du narrateur (par une analyse des intensificateurs, du discours direct et des questions ou exclamations vers l'expérimentateur).

Ainsi, notre analyse des évaluations sociales montre que, conformément à notre hypothèse, les récits des patients cérébrolésés comportent davantage d'intensificateurs et d'énoncés de discours direct que ceux des sujets contrôles. Ces résultats sont à mettre en perspective avec les données de la littérature concernant les déficits du contrôle émotionnel souvent décrits chez les patients cérébrolésés, en particulier chez les patients porteurs de lésions frontales (Blumer et Benson, 1975 ; Eslinger, Flaherty-Craig et Benton, 2004). En effet, selon nous, des déficits dans la régulation des émotions (désinhibition émotionnelle et comportementale) peuvent avoir des répercussions sur les narrations des patients cérébrolésés, car la production d'un récit socialement adapté nécessite un contrôle émotionnel qui apparaît difficile à maintenir chez ces patients.

Dans le futur, les données recueillies auprès d'une population plus vaste nous permettront d'envisager de nouvelles possibilités en terme de remédiation. Enfin, la constitution d'échantillons homogènes (en terme de localisation, d'âge d'acquisition des lésions et de passation, de latéralisation et d'étendue des lésions) sera également nécessaire afin de mieux

comprendre le rôle des différentes régions cérébrales impliquées dans le domaine de la pragmatique et afin d'évaluer d'éventuels phénomènes de plasticité cérébrale.

Références

Références bibliographiques

- BELLUGI, Ursula, REILLY, Judy, KRIETER, Jacqueline, DOYLE, Teresa & JONES, Wendy (2003). Language and non linguistic cognition in children with Williams syndrome: A complex interaction, *Enfance*, 3: 237-249.
- BERMAN, Ruth & SLOBIN, Dan (1994). *Relating events in a narrative*. Hillsdale, NJ : Lawrence Erlbaum Associates.
- BERNICOT, Josie & DARDIER, Virginie (2001). Les troubles de la communication consécutifs aux lésions frontales: l'exemple de la situation d'interview, *Revue de Neuropsychologie*, 10: 281-331.
- BERNICOT, Josie, DARDIER, Virginie, FAYADA, Catherine, PILLON, Bernard & DUBOIS, Bruno (2001). Pragmatique et métapragmatique chez deux patients souffrant de lésions frontales unilatérales, *Psychologie de l'Interaction*, 13-14: 141-184.
- BERNICOT, Josie, LACROIX, Agnès & REILLY, Judy (2003). La narration chez les enfants atteints du syndrome de Williams : aspects structuraux et pragmatiques, *Enfance*, 3: 265-283.
- BIDDLE, Kathleen, MCCABE, Allyssa & BLISS, Lynn (1996). Narrative skills following traumatic brain injury in children and adults, *Journal of Communicative Disorders*, 29: 447-469.
- BLUMER, David & BENSON, Dietrich (1975). Personality changes with frontal and temporal lobe lesions, in *Psychiatric Aspects of Neurologic Disease*, Benson, David & Blumer Dietrich (Eds). New-York : Grune & Stratton, 151-170.
- CHAMPAGNE, Maud, VIRBEL, Jacques, NESPOULOUS, Jean-Luc & JOANETTE, Yves (2003). Impact of right hemispheric damage on a hierarchy of complexity evidenced in young normal subjects, *Brain and Cognition*, 53: 152-157.
- CHAPMAN, Sandra, LEVIN, Harvey, WANEK, Alicia, WEYRAUCH, Julie & KUFERA, Joseph (1998). Discourse after closed head injury in young children, *Brain and Language*, 61: 420-449.
- COELHO, Carl, LILES, Betty & DUFFY, Robert (1995). Impairments of Discourse Abilities and Executive Functions in Traumatically Brain Injured Adults, *Brain Injury*, 9: 471-477.
- DARDIER, Virginie (2001). *Pragmatique du langage et de la communication chez l'adulte et chez l'enfant : une étude de patients porteurs de lésions frontales*. Thèse nouveau régime, Paris, Université Paris V, manuscrit non publié.
- DARDIER, Virginie (2004). *Pragmatique et pathologies – Comment étudier les troubles de l'usage du langage*. Paris : Bréal.
- DARDIER, Virginie & BERNICOT, Josie (2000). Les troubles de la communication consécutifs aux lésions frontales : l'exemple de la situation d'interview, *Revue de Neuropsychologie*, 10, 2: 281-331.
- DARDIER, Virginie, DELAYE, Corinne & LAURENT-VANNIER, Anne (2003). La compréhension des actes de langage par des enfants et des adolescents porteurs de lésions frontales : l'analyse des demandes, *Enfance*, 3: 223-236.
- DARDIER, Virginie, REILLY, Judy, BATES, Elizabeth, DELAYE, Corinne & LAURENT-VANNIER, Anne (2004). La cohésion du discours chez les enfants et les adolescents cérébrolésés – Analyse de l'usage des pronoms et des conjonctions, *Cahiers d'Acquisition et de Pathologie du Langage*, 24: 101-114.
- DENNIS, Maureen, JACENNIK, Barbara & BARNES, Marcia (1994). The content of narrative discourse in children and adolescents after early-onset hydrocephalus and in normally-developing age peers, *Brain and Language*, 46: 129-165.

- ESLINGER, Paul, FLAHERTY-CRAIG, Claire & BENTON, Arthur (2004). Developmental outcomes after early prefrontal cortex damage, *Brain and Cognition*, 55: 84-103.
- GLOSSER, Guila & DESER, Tony (1990). Patterns of discourse production among neurological patients with fluent language disorders, *Brain and Language*, 40: 67-88.
- JOANETTE, Yves, GOULET, Pierre & HANNEQUIN, Didier (1990). *Right hemisphere and verbal communication*. New-York : Springer-Verlag.
- KERN, Sophie (1997). *Comment les enfants jonglent avec les contraintes communicationnelles, discursives et linguistiques dans la production d'une narration*. Lille : Presses du Septentrion.
- KERN, Sophie (2002). Contexte et acquisition du langage : un cas de narration. *French Language Studies*, 12, 189 – 209.
- LILES, Betty, COELHO, Carl, DUFFY, Robert & ZALAGENS, Mary (1989). Effects of elicitation procedures on the narratives of normal and closed-head injured adults, *Journal of Speech & Hearing Research*, 54: 356-366.
- LOSH, Molly, BELLUGI, Ursula, REILLY, Judy & ANDERSON, Diane (2000). Narratives as a social engagement tool: The excessive use of evaluation in narratives from children with Williams syndrome, *Narrative Inquiry*, 10: 265-290.
- MCDONALD, Skye & PEARCE, Samantha (1998). Requests that overcome listener reluctance: Impairment associated with executive dysfunction in brain injury, *Brain and Language*, 61: 88-104.
- MAR, Raymond (2004). The neuropsychology of narrative: story comprehension, story production and their interrelation, *Neuropsychologia*, 42: 1414-1434.
- MAYER, Mercer (1969). *Frog, where are you?* New-York : Dial Press.
- MENTIS, Michele & PRUTTING, Carol (1991). Analysis of topic as illustrated in a head-injured and a normal adult, *Journal of Speech and Hearing Research*, 34: 583-595.
- PETER-FAVRE, Claire (1995). Conversations avec une patiente souffrant de lésions traumatiques bifrontales : Ajustements mutuels, *Revue de Neuropsychologie*, 5: 53-85.
- PETER-FAVRE, Claire (1999). *Communication verbale et pathologie post-traumatique : Interactions conversationnelles avec une patiente souffrant de lésions bifrontales*. Thèse de Doctorat, Université de Nancy II.
- PETER-FAVRE, Claire (2002). Discours et conversation chez les traumatisés crânio-cérébraux, in *Pragmatique et psychologie*, Bernicot, Josie, Trognon, Alain, Guidetti, Michèle & Musiol, Michel (Eds). Nancy : Presses Universitaires de Nancy, 281-304.
- REILLY, Judy, BATES, Elizabeth & MARCHMAN, Virginia (1998). Narrative discourse in children with early focal brain injury, *Brain and Language*, 61: 335-375.
- REILLY, Judy, LOSH, Molly, BELLUGI, Ursula & WULFECK, Beverly (2004). Frog, where are you? Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome, *Brain and Language*, 88: 229-248.
- REILLY, Judy & WULFECK, Beverly (2004). Issues in plasticity and development: Language in atypical children, *Brain and Language*, 88: 163-166.
- STEMMER, Brigitte & COHEN, Henri (2002). Neuropragmatique et lésions de l'hémisphère droit, *Psychologie de l'Interaction*, 13-14: 15-46.

FORMES LINGUISTIQUES ADJECTIVALES ET IDENTIFICATION DES REPRÉSENTATIONS COGNITIVES

Gaëlle DELEPAUT

g_delepaut@yahoo.fr

LCPE, 11, rue de Lourmel

75015 PARIS

Abstract : This study concerns the identification of semantic properties of comfort during railroad trips. A linguistic analysis of adjectival affixes and the cognitive inferences they entail was carried out. Their morphology contributes to the identification of different kinds of representations of comfort in trains when asking passengers about their expectations, their feelings concerning the journey and their ideal representations.

1. Introduction

L'analyse linguistique que nous proposons s'inscrit en réponse à une demande de la SNCF relative à la perception du confort global en train. Les objectifs du projet sont d'identifier les propriétés sémantiques du confort en train et leurs relations de dépendance, à partir d'analyses linguistiques et cognitives. Cet article porte sur la manière dont les formes linguistiques en contexte, utilisant les ressources de la langue mises en discours, renseignent sur les structures cognitives construites à partir des perceptions sensorielles. On tente ici d'identifier, à partir de l'analyse des formes adjectivales, les représentations individuelles et partagées qui se construisent dans les discours des voyageurs, lorsqu'on les interroge sur leur expérience sensible du confort.

1.1. Hypothèses

L'analyse s'appuie sur des hypothèses psychologiques relatives à l'organisation des connaissances en mémoire (Rosch, 1978). Les connaissances sont structurées en catégories cognitives à partir de l'expérience sensible individuelle (Dubois, 2000). Ces catégories relèvent de la perception sensorielle dans différentes modalités sensibles. Elles peuvent être caractérisées par des propriétés descriptives. Le partage de propriétés entre plusieurs concepts permet d'identifier des « ressemblances de famille » et des représentations plus ou moins typiques au sein de catégories aux frontières floues. La première question que l'on se pose alors est la suivante : le confort en train est-il une catégorie cognitive de ce type ?

Notre hypothèse linguistique porte sur la régularité des formes et des constructions langagières, qui permet d'identifier des catégories linguistiques. On s'intéresse en particulier, dans cet article, aux travaux sur la morphologie des adjectifs (Dubois et Dubois-Charlier, 1999). Les adjectifs construits sur des verbes (ou déverbaux, cf. Henrichsen, 1967 : 97) sont étudiés en regard de la sémantique :

- du verbe qui peut exprimer par exemple, l'évaluation comme dans « gênant » (→ « gêner »), ou la fonctionnalité, comme dans « pliable » (→ « plier ») ;
- de l'affixe (Bybee, 1988). Trois affixes sont retenus :
 - -ant, est parfois paraphrasé par une relative, comme dans « gênant » → « qui gêne » ou « qui me gêne », selon les théories (David *et al.*, 1997) ;
 - -é (et allomorphes -i, -u, -is, -it) est généralement associé à « qui est + verbe », comme dans : « détendu » → « qui est détendu » ;

- -ble (soit -able, soit l'allomorphe -ible), que l'on peut considérer comme la possibilité (« portable » → « que je peux porter »), l'expression d'un effet sur le sujet associé au plaisir ou déplaisir (« appréciable », David *et al.*, 1997 : 66-67) ou « le procès réalisé comme caractéristique de l'entité qualifiée » (Corbin, 1987 : 375-376).

L'hypothèse qui est alors formulée porte sur la signification des affixes des adjectifs déverbaux en tant qu'indices d'un jugement, ou d'un état du sujet, ou d'un effet sur le sujet.

Le couplage des hypothèses cognitives et linguistiques conduit à postuler l'existence de liens complexes entre langage et fonctionnement cognitif dans l'expression de l'expérience sensorielle. Nous formulons l'hypothèse que certaines formes linguistiques exprimées en discours sont privilégiées pour l'identification des représentations cognitives (David, 1997). En particulier, on développera la contribution d'une analyse linguistique de la morphologie des adjectifs déverbaux pour l'identification de l'organisation cognitive des propriétés sensorielles du confort en train. Les résultats de cette analyse seront interprétés dans le cadre spécifique des contextes discursif et de questionnement (ordre des questions).

1.2. Méthodologie

La méthodologie d'enquête employée procède par questionnaires ouverts écrits. Ce choix est orienté :

- d'une part, par les objectifs de l'enquête qui portent, en psychologie, sur la perception du monde sensible, et
- d'autre part, par l'analyse envisagée qui se situe en linguistique cognitive.

L'intérêt du questionnaire ouvert est de ne pas orienter les réponses à partir d'*a priori* physicalistes (Mzali, 2002) ; il vise ainsi à accéder au ressenti individuel des voyageurs. La démarche est d'abord fondée en sciences humaines et sociales à travers des modes de questionnements spécifiques (Dubois, 1996) et des analyses des constructions discursives produites dans les réponses. L'intérêt des analyses linguistiques est de s'intéresser, non pas seulement au contenu des données des informateurs comme le font les sociologues, mais également au fonctionnement discursif et sémantique des énoncés, et donc à l'expression en discours de représentations individuelles (par l'expérience) et partagées (par l'usage d'une langue commune).

Le questionnaire comporte quinze questions qui s'organisent en axes principaux, du plus général au plus spécifique, et qui conditionnent la formulation et l'ordre des questions. Les axes de questionnement structurent les questions successivement sur : les représentations en mémoire, le ressenti en cours, et des représentations idéales. 240 voyageurs ont été interrogés, également distribués :

- dans 4 TGV (les noms des différents matériels sont : 1) la Salle Haute du Duplex, 2) la Salle Basse du Duplex, 3) le PSE et 4) le Réseau¹ ; ils seront respectivement abrégés SH, SB, SE et R dans les graphiques suivants) et 2 Corails (dont les noms sont 1) Corail et 2) Téoz, abrégés C et T),
- en première et en seconde classe (pour les différencier, les nombres 1 et 2 sont associés aux lettres des matériels ferroviaires dans les graphiques).

2. Analyse du corpus

2.1. La morphologie des adjectifs

Une question à laquelle nous tentons de répondre concerne la validité des propriétés morphologiques comme indicateurs des propriétés sémantiques du confort en train. Tous les

¹ Ces deux derniers TGV sont des TGV à un niveau.

énoncés du corpus ont fait l'objet d'une première analyse visant à identifier les adjectifs dans leur contexte discursif. Certains énoncés n'ont pas permis une sélection définitive, car les catégories linguistiques peuvent être ambiguës ou indéterminées. En effet, l'énoncé « *rien à cause des wagons fumeurs que j'apprécie tout particulièrement* », fourni en réponse à la question 09 « *que manque-t-il à votre confort dans ce train lorsque vous êtes assis(e) ?* », illustre, avec le terme « *fumeurs* » la difficulté à reconnaître les catégories linguistiques (voir Noailly, 1990). De même, dans l'énoncé « *un coussin appui-tête* », si, dans tous les cas canoniques, le terme « *appui-tête* » est considéré comme un substantif, il est ici, en discours, employé comme un adjectif et permet de poser la question de la stabilité des catégories linguistiques indépendantes des contextes discursifs. Entre autres cas, on relève également la différenciation entre des emplois adjectivaux et verbaux, qui peut être opérée à partir d'un test de quantification par l'expression du degré absolu représenté par l'adverbe « très ».

Au sein des adjectifs, on repère les formes construites et non construites. Par exemple, dans l'énoncé « *trajet **rapide**, sans retard, bien **climatisé*** », donné en réponse à la question 04 « *quelles sont vos attentes pour ce voyage en train ?* », « *climatisé* » est construit et « *rapide* » est non construit. Les adjectifs construits peuvent avoir pour base différentes catégories grammaticales. Les adjectifs désadjectivaux ne nous intéressent pas ici, car ils sont rares dans le corpus recueilli et ne qualifient pas des propriétés du confort en train. Par contre, on s'intéressera aux formes dénominales : par exemple, dans l'énoncé « *une ouverture **automatique** des portes entre les voitures* » fourni en réponse à la question 09, l'adjectif est construit sur un substantif (« automate »). Ces adjectifs réfèrent à une source et leur intérêt pour cette recherche réside dans le fait qu'ils renvoient, par le biais d'un qualificatif, à une certaine objectivation de l'objet désigné. Les adjectifs déverbaux peuvent s'appliquer au monde ou au sujet, et ainsi être indices de propriétés du confort en train ou d'évaluations de la part des voyageurs, par exemple : « *un service mini-bar **ambulante** serait très **appréciable** / calme / ponctualité du train* ».

Les suffixes de ces adjectifs déverbaux sont nombreux et nous nous bornerons aux trois présentés dans le cadre théorique². Nous vérifierons les hypothèses selon lesquelles :

- les adjectifs déverbaux en -ble correspondent à l'expression d'une possibilité ;
- les adjectifs déverbaux en -é seraient relatifs à des événements passés, parce qu'ils sont généralement issus de participes passés ;
- les adjectifs déverbaux en -ant renverraient à des événements présents, parce qu'ils sont massivement issus de participes présents.

Il est désormais envisageable d'appliquer ces « outils linguistiques » à l'analyse de trois questions issues de chacun des axes du questionnaire afin de vérifier les hypothèses formulées et la validité de ces indicateurs de représentations cognitives :

- sur les attentes des voyageurs avant le trajet au cours duquel ils sont interrogés : la question 04 « *quelles sont vos attentes pour ce voyage en train ?* » ;
- sur le trajet en cours : la question 09a « *pour vous, que manque-t-il à votre confort dans ce train lorsque vous êtes assis(e) ?* » ;
- sur les trajets à venir : la question 14 « *pour vous, comment serait le train idéal ?* ».

L'analyse proposée est une analyse qualitative, qui s'appuie, pour illustrer notre propos, sur des données chiffrées, issues d'un grand corpus.

² C'est pourquoi des écarts entre le nombre total d'adjectifs déverbaux et le détail des suffixes de ces adjectifs peuvent être observés dans l'analyse.

2.2. Les représentations du trajet en train

Une première observation vise à montrer la répartition des adjectifs selon les questions, tout en distinguant à cette étape les adjectifs construits et les non-construits :

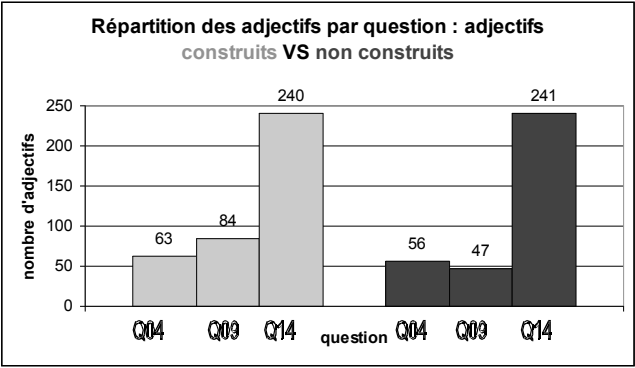


Figure 1 - Les adjectifs construits et non construits

Cet aperçu global permet déjà de mettre en avant la productivité de la question 14 par rapport aux autres questions étudiées (et du questionnaire). En revanche, globalement, on n’observe pas d’écart significatif dans l’emploi des adjectifs construits par rapport aux non-construits.

2.2.1. Les trajets antérieurs

La question sur les trajets antérieurs (question 04) visait à identifier les représentations en mémoire. L’utilisation des « outils linguistiques » présentés dans les parties précédentes à la question sur les attentes participe à ce travail. Ainsi, dans l’ensemble des matériels ferroviaires et des classes, sur les 119 adjectifs relevés 63 sont construits :

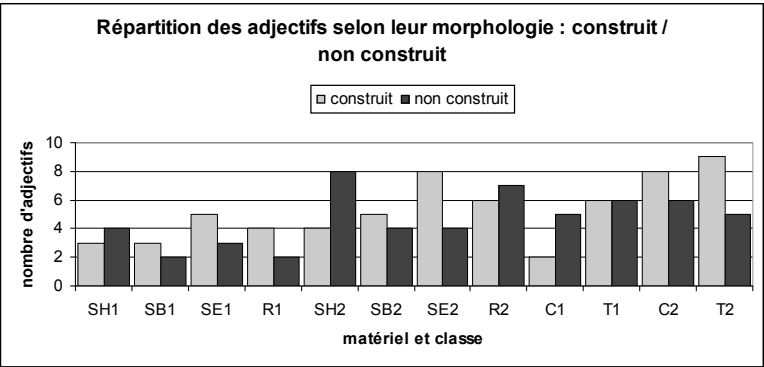


Figure 2 - Les adjectifs construits et non construits de la question 04

L’analyse de ces adjectifs en contexte permet de mettre en avant la rareté des adjectifs qualifiant des substantifs dans tous les trains étudiés. On observe en revanche des disparités en 2nde classe : le TGV PSE est le seul permettant de recueillir deux fois plus d’adjectifs construits que non construits, et dans le Téo2 9 sur 14 sont construits.

La construction des adjectifs, selon les catégories morphologiques étudiées, est décrite Figure 3 :

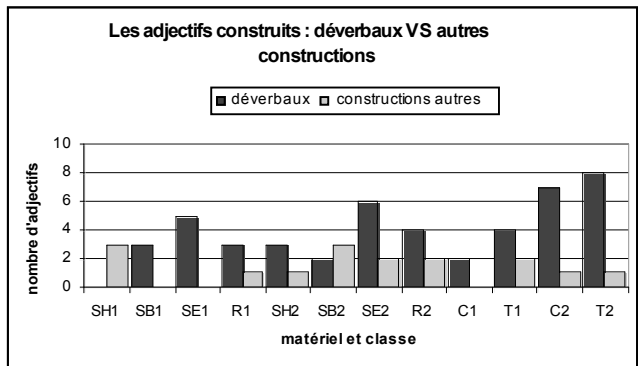
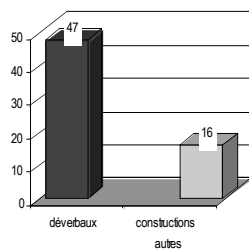


Figure 3 - Les adjectifs construits et non construits de la question 04

Globalement, plus d'adjectifs déverbaux (47) que d'autres constructions (16) sont relevés. Les voyageurs réfèrent donc davantage à leur ressenti. Des différences inter-matériels peuvent aussi être soulignées : en particulier, la Salle Haute du TGV Duplex en 1^{ère} classe ne comporte aucun adjectif déverbal, et en contraste le TéoZ en 2^{nde} classe est particulièrement productif de ce genre de formes. La répartition des adjectifs déverbaux selon les constructions étudiées est fournie dans la Figure 4 ci-dessous :

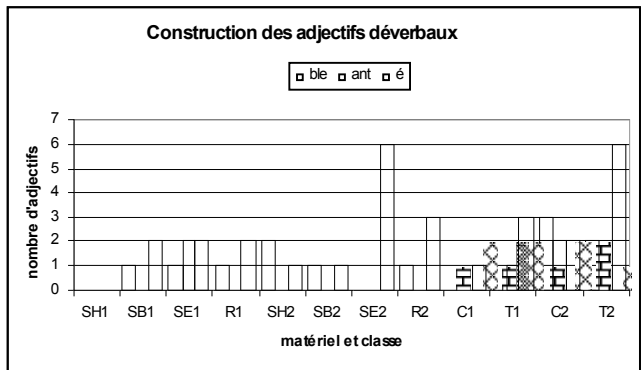
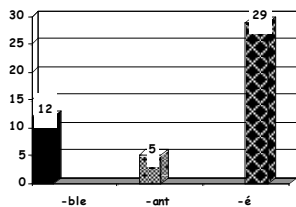


Figure 4 - Les adjectifs déverbaux de la question 04

Ainsi, on observe que la question sur les trajets antérieurs ne suscite qu'exceptionnellement des adjectifs avec le suffixe -ant (5 sur les 47 adjectifs déverbaux relevés parmi les 119 adjectifs de cet extrait de corpus). Les adjectifs construits sur -é concernent le sujet (10), puis le trajet (7) et enfin l'objet - le train - (6). Ces résultats confirment donc l'hypothèse selon laquelle la morphologie différentielle des adjectifs déverbaux permet d'identifier que les attentes des voyageurs sont basées sur les représentations en mémoire. Sur la base de ces résultats, nous allons maintenant voir s'il est possible de les généraliser.

2.2.2. Le trajet en cours

La question sur le trajet en cours (question 09) visait à effectuer un état des lieux du confort du train étudié. Globalement, la proportion d'adjectifs construits est très supérieure (84) à celle des adjectifs non construits (47). De plus, la répartition par matériel indique l'emploi massif d'adjectifs à bord du Corail (20 en 2^{nde} classe et 19 en 1^{ère}), par rapport aux autres situations. En outre, le Corail en 2^{nde} classe compte plus d'adjectifs non construits que de construits (respectivement 11 et 9), alors que 14 adjectifs sont construits sur les 19 adjectifs relevés en Corail 1^{ère}. Si les adjectifs construits renvoient davantage à des discours exprimant la subjectivité des voyageurs, on observe ici davantage de discours inscrits dans un style

convenu en 2nde classe (34 adjectifs construits VS 29 non construits), qu'en 1^{ère} classe (50 adjectifs construits VS 24 adjectifs non construits) où les discours s'inscrivent dans l'expression de l'expérience individuelle.

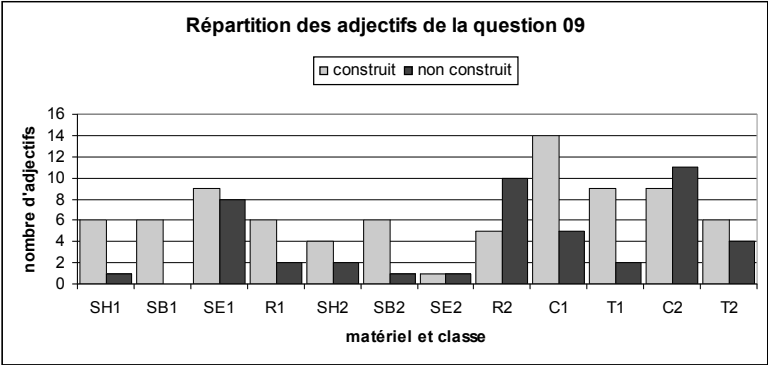


Figure 5 - Les adjectifs construits et non construits de la question 09

Sur les 84 adjectifs construits des réponses à la question 09, 44 sont déverbaux. La répartition selon le type de matériel ferroviaire est instructive en 1^{ère} classe, en particulier en Corail, où les discours comportant des déverbaux concerneraient des ressentis, alors que ceux du TGV PSE feraient référence à des sources avec un emploi massif d'adjectifs dénominaux. La construction des adjectifs déverbaux est observable dans la Figure 6 :

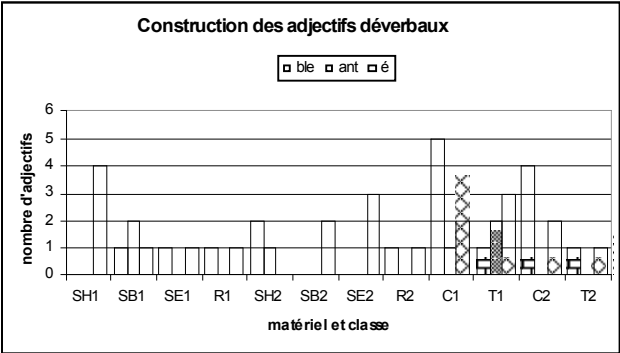
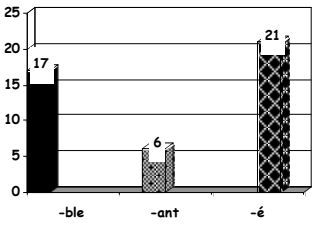


Figure 6 - Les adjectifs déverbaux de la question 09

Globalement plus d'adjectifs déverbaux suffixés en -é sont employés. Cependant, on observe l'emploi plus important d'adjectifs en -ble dans certaines situations, en particulier en Corail : on identifie ainsi des références au confort variées.

L'abondance de ces constructions déverbaux met en avant la rareté du suffixe -ant pour la description d'une expérience en cours. Ce résultat n'est pas contradictoire avec ceux développés en 2.2.1, mais participe à la formulation d'interprétations complémentaires et plus précises. En effet, la morphologie des adjectifs en discours et en contexte indique que le confort perçu correspond pour le suffixe -é, à un état :

- du sujet, par exemple : « moins *secoué* - pas possible d'écrire correctement donc plus de stabilité, moins de bruit » ;
- de l'objet : « fauteuil trop *incliné* ».

Ce résultat est étayé par l'analyse des adjectifs en -ble se rapportant à la fonctionnalité des éléments, si l'on considère à la fois la sémantique du verbe et celle du suffixe, par exemple : « une petite tablette, un coin discret, des fauteuils *inclinables* » (verbe « incliner » et suffixe

« -ble »). Ainsi, le questionnement conditionne la formulation des réponses et l'analyse qui en est faite se doit d'en tenir compte. En outre, on valide l'hypothèse selon laquelle la morphologie contribue à l'identification des propriétés sémantiques : en effet, le suffixe -ble est un des indices principaux d'une propriété du confort mentionnée par les voyageurs et relative à une catégorie sémantique que nous avons désignée « contraintes et possibles »³.

2.2.3. Les trajets à venir

Concernant les trajets à venir par l'évocation du train idéal (question 14), la productivité de cette question est inattendue, d'autant qu'elle intervient en fin de questionnaire. On y relève 240 adjectifs construits et 241 non construits. La moitié des adjectifs construits correspond à des déverbaux ; parmi les 120 constructions autres, l'emploi massif de dénominiaux (« *silencieux* », par exemple) est l'indice, fréquent mais non systématique, de renvois à des objets perceptibles. La construction des adjectifs déverbaux est détaillée Figure 7 :

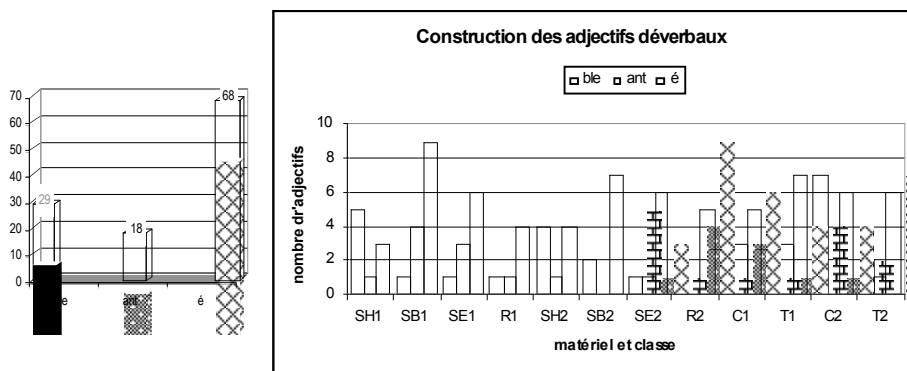


Figure 7 - Les adjectifs déverbaux de la question 14

La morphologie des adjectifs déverbaux permet d'induire qu'il n'y a pas de description de l'expérience en cours comme ressentie et évaluée (emploi de seulement 18 adjectifs suffixés par -ant), mais comme productrice d'effets et décrivant des états des objets idéaux.

En effet, le contexte des adjectifs relevés - construits ou non - indique que les représentations idéales :

- se construisent sur des possibles (avec les adjectifs suffixés par -ble), par exemple : « *systématiquement : un bar restaurant, stabilité, repose-pied **escamotable*** » ;
 - sont influencées par les représentations en mémoire et en cours d'expérience : « *un train très **confortable** (toilettes **propres**) avec un service **minimum assuré** lors de grèves **justifiées**, arrivant à l'heure* » ;
 - relèvent des états des objets ou du train lui-même : « ***convivial** (café) **climatisé** - bien **chauffé insonorisé** (pas de portable) non fumeur fauteuil **large et réglable*** ».
- On observe ici la sémantique du suffixe -isé qui induit un processus technique.

3. Conclusions

En conclusion, nous avons pu montrer que les discours obtenus en réponse au questionnaire permettent d'inférer une certaine représentation du monde identifiable par l'analyse linguistique des énoncés qui permet l'accès au ressenti individuel et aux représentations

³ Par exemple, l'énoncé « *espace vitré, vitesse, bar **disponible*** » fourni en réponse à la question « *selon vous, en quoi votre trajet est confortable ?* », illustre précisément une évaluation où l'énonciateur n'indique pas s'il est allé ou compte se rendre au bar et si celui-ci est confortable ou non, mais que le simple fait de 'pouvoir en disposer' contribue fortement à son confort.

cognitives. Les formes discursives employées pour l'expression des différents jugements montrent l'intérêt d'analyses linguistiques successives pour un même corpus. L'analyse effectuée porte dans sa totalité sur plusieurs phénomènes complémentaires issus d'observations faites en syntaxe, en morphologie et en sémantique lexicale et appuyées sur des hypothèses cognitives. La démarche procède ainsi par analyses successives entre formes linguistiques et hypothèses cognitives pour formuler, étayer et valider des hypothèses, voire généraliser les inférences réalisées ou bien les préciser en restreignant la validité de certaines interprétations. Ici en particulier, la morphologie des adjectifs, notamment celle des adjectifs déverbaux, permet d'identifier différentes représentations cognitives relatives au confort en train. Les possibilités qu'offre la langue aux voyageurs permettent ainsi de situer leurs perceptions dans un jugement évaluatif et dans des expériences vécues. Car, si on ne peut affirmer que les termes employés renvoient de manière consciente aux différents éléments qui les composent, leur sémantique et le choix de certaines formes parmi celles disponibles en langue contribuent cependant à la cohérence des énoncés et à l'expression de propriétés du confort en train. Ainsi, le suffixe -ble a majoritairement une valeur évaluative ou des possibles exprimés pour les représentations en cours et idéales. Le suffixe -é exprime des états du sujet ou de l'objet. Le suffixe -ant exprime le plus souvent des évaluations en cours. Ces résultats ne s'inscrivent pas dans le cadre d'une sémantique régulière des suffixes et ne sont valables qu'à condition d'inscrire l'analyse morphologique et les différentes interprétations d'un même élément morphologique dans le contexte du questionnaire et en discours où se réalise la sémantique des suffixes. Le recoupement des différents domaines abordés pour les analyses participe à l'identification de l'implication des voyageurs dans leur discours et au développement des connaissances sur les liens entre langage et cognition.

Références

Références bibliographiques

- BYBEE, Joan L. (1988). Morphology as lexical Organisation, in *Theoretical Morphology*, Hammond, Michael & Noonan, Michale (Eds). San Diego / N.Y.: Academic Press, 119-141.
- CORBIN, Danièle (1987). *Morphologie dérivationnelle et structuration du lexique*. Thèse de Doctorat d'Etat de l'Université de Paris VIII, Tübingen : Max Niemeyer Verlag.
- DAVID, Sophie (1997). Représentations sensorielles et marques de la personne : contraste entre olfaction et audition, in *Catégorisation et cognition*, Dubois, Danièle (Ed). Paris: Kimé, 211-242.
- DAVID, Sophie, DUBOIS, Danièle, ROUBY, Catherine & SCHAAL, Benoît (1997). L'expression des odeurs en français : analyse lexicale et représentation cognitive, *Intellectica*, 24: 51-83.
- DUBOIS, Danièle (2000). Categories as Act of Meaning, *Cognitive Science Cognitive Quarterly*, 1, Hermes Science Publications, 35-68.
- DUBOIS, Danièle (1996). Matériels et consignes : un type de questionnement social dans la recherche expérimentale en linguistique, *Le questionnement social*, n° spécial des *Cahiers de Linguistique Sociale*: 89-98.
- DUBOIS, Jean & DUBOIS-CHARLIER, Françoise (1999). *La dérivation suffixale en français*. Paris : Nathan Université.
- HENRICHSEN, Arne-Johan (1967). Quelques remarques sur l'emploi des formes verbales en -ant en français moderne, *Revue Romane*, 2: 97-107.
- MZALI, Myriam (2002). *Perception de l'ambiance sonore et évaluation du confort acoustique dans les trains*. Thèse de doctorat de l'Université de Paris VI, Paris.
- NOAILLY, Michèle (1990). *Le substantif épithète*. Paris : PUF.
- ROSCH, Eleanor (1978). Principles of Categorization, in *Cognition and Categorization*, Rosch, Eleanor & Lloyd, Barbara (Eds). New-York: Hillsdale, N.J., 27-48.

"UN OURS" / "DES NOURS" OU LE RÔLE DE LA FRÉQUENCE SUR L'ACQUISITION DE LA LIAISON EN FRANÇAIS¹

Céline DUGUA

Celine.Dugua@u-grenoble3.fr

Laboratoire Lidilem – Université Stendhal

38040 Grenoble

Abstract : The French liaison is a phonological alternation phenomenon, whose acquisition constitutes a difficulty for children approximately until the age of 5. Our experimental studies allowed us to identify several stages in this acquisition. We argue that this phonological development is guided by usage, in particular by frequency effects.

1. Introduction

La liaison est un cas d'alternance phonologique qui consiste en l'apparition d'une consonne entre deux mots. Particularité linguistique largement étudiée depuis les années 1960, elle a servi de validation aux différentes théories phonologiques qui ont eu cours (phonologie générative, phonologie multilinéaire, théorie de l'optimalité, théorie des constructions). Toutefois, ce phénomène ne se limite pas à la phonologie et fait, au contraire, intervenir d'autres niveaux linguistiques tels que la syntaxe, la morphologie et la pragmatique. Quant à la question de son acquisition en français, elle est restée longtemps ignorée jusqu'aux premières recherches systématiques menées par Wauquier-Gravelines (2002; 2003) et Chevrot et Fayol (2000). Dans cet article, nous nous intéresserons à un aspect particulier influençant cette acquisition, à savoir le rôle de la fréquence. En effet, nous soutenons, avec les tenants des théories basées sur l'usage (Bybee, 2003 [2001]; Langacker, 1987; Tomasello, 2003), que la fréquence d'apparition d'un élément ou d'une séquence dans l'input, influence sa disponibilité dans le lexique et par là même les processus en jeu dans son acquisition.

Nous commencerons par préciser ce que revêt le phénomène de liaison et les raisons pour lesquelles il constitue une difficulté pour les enfants. Ensuite, nous présenterons le scénario développemental élaboré à partir des études sur la liaison. Enfin, nous nous attacherons aux hypothèses, méthodologies et résultats qui ont guidé cette recherche sur le rôle qu'exerce la fréquence sur l'acquisition de la liaison.

2. La liaison et les difficultés qu'elle pose aux enfants

La liaison consiste en l'apparition d'une consonne (/n/, /z/ ou /t/ dans 99.8 % des cas, Boë et Tubach, 1992)² entre deux mots – un Mot1 et un Mot2 – qui présentent un lien syntaxique fort. En outre, pour qu'une liaison apparaisse, le Mot2 doit commencer par un son vocalique lorsqu'il est produit en isolation par un adulte : on trouve, par exemple, une liaison devant *arbre* [arbr] dans *un arbre*, mais pas devant *camion* [kamjɔ̃] dans *un camion*. Par ailleurs, la consonne de liaison (désormais CL) n'est pas produite à la fin du Mot1 lorsqu'il se trouve en finale d'énoncé (*j'en prends un* prononcé [ʒãprãœ] sans consonne finale). Cette CL n'est

¹ Cette recherche est financée par l'ACI Systèmes Complexes (2004-2007) : *Pati Papa ? Modélisation de l'émergence d'un langage articulé dans une société d'agents sensori-moteurs en interaction.*

² Ce résultat a été obtenu à partir du tableau présenté par Boë et Tubach (1992 : 169).

jamais prononcée par l'adulte à l'initiale du Mot2 lorsqu'il se trouve en début d'énoncé : par exemple, la liaison en /z/ est réalisée dans *les ânes*, en revanche, elle n'apparaît pas dans l'apostrophe *âne*. Enfin, la nature phonétique de la CL – /n/, /z/ ou /t/ – dépend du Mot1 : les Mots1 tels que *un* ou *aucun* déclenchent une liaison en /n/, les Mots1 *petit*, *grand*, etc. une liaison en /t/ et les Mots1 *gros*, *très*, *deux*, etc. une liaison en /z/.

La liaison est un phénomène linguistique long à acquérir, essentiellement pour deux raisons. Tout d'abord, l'apparition de la consonne de liaison devant un Mot2 à initiale vocalique entraîne une resyllabification. Dans la séquence *un arbre*, le découpage syllabique /œ.naʁbʁ/ fait apparaître une frontière devant la forme /naʁbʁ/. En d'autres termes, en contexte de liaison, les frontières de syllabes et de mots ne sont plus concordantes, ce qui complexifie la tâche de segmentation pour les enfants. La seconde difficulté inhérente à ce phénomène est liée à la nature variable de la CL en fonction du Mot1. En confrontant cette caractéristique à la précédente, on peut penser que l'enfant récupèrera des formes lexicales différentes en fonction du contexte dans lequel elles ont été entendues. Par exemple, il récupèrerait /naʁbʁ/ dans *un arbre*, /zaʁbʁ/ dans *les arbres*, /taʁbʁ/ dans *grand arbre*.

Suite aux premières esquisses de scénario développemental proposées par Morin (2003 [1998]), qui rendaient compte du lien entre les erreurs en contexte de liaison et la façon dont les enfants font face aux difficultés de segmentation et de formation des mots, nous avons étayé cette première version du scénario en mettant en place plusieurs démarches empiriques permettant d'observer la mise en place de la liaison et les phénomènes qui lui sont liés (Chevrot *et al.*, 2005).

3. Scénario développemental de l'acquisition de la liaison

Le scénario développemental dont nous allons faire état a été établi à partir de l'analyse de productions d'enfants âgés de 2 à 6 ans, en situations naturelle et expérimentale (Dugua, 2005).

3.1. Ancrage théorique

Nous interpréterons nos données au regard des principes des théories basées sur l'usage. Dans cette approche, l'usage du langage est la clé pour comprendre à la fois sa genèse, son fonctionnement et son évolution. Afin de rendre compte de l'impact de l'usage, les recherches intégrées à ces approches tentent notamment de mettre à jour le rôle de la fréquence d'apparition d'un type de séquences particulier dans l'input sur les productions enfantines précoces (Cameron-Faulkner *et al.*, 2003; Lieven *et al.*, 2003).

Par ailleurs, un second principe, emprunté aux grammaires de constructions et auquel les théories basées sur l'usage font expressément référence, s'appuie sur l'idée que l'unité de base du traitement linguistique est la construction. Les constructions sont des paires forme/sens conventionnalisées (Goldberg, 1995), c'est-à-dire que dans chaque construction une forme phonologique/syntaxique correspond à une intention sémantique/pragmatique. En outre, les constructions peuvent être de formes diverses, elles varient en effet selon deux dimensions principales (Diessel, 2004). D'une part, elles se distinguent en termes de complexité syntagmatique ; elles peuvent en effet correspondre à un seul mot, à des séquences telles que *déterminant + nom*, *clitique + verbe*, à des structures syntaxiques telles que le passif, etc. D'autre part, les constructions varient selon une échelle de schématisation et d'abstraction. En d'autres termes, on trouverait dans le lexique des constructions concrètes, de type expressions idiomatiques, mots, ou fragments d'énoncés récupérés dans l'input et stockés tels quels (par exemple : *tout à coup*, *mon doudou*, *caché*, etc.) et des constructions abstraites et schématiques appelées *schémas* qui peuvent intégrer des catégories grammaticales (groupe nominal, verbe, etc.) correspondant à des emplacements

libres, ou *slots* (par exemple : *j'ai + Vê*). Les schémas abstraits émergeraient à partir de processus cognitifs (généralisation, schématisation, analogie, voir Langacker, 1987; Tomasello, 2003) appliqués aux premières constructions concrètes mémorisées. En ce sens, les constructions concrètes et abstraites sont liées entre elles puisque les secondes ne peuvent émerger qu'à partir des premières.

À partir de la présentation de cette unité particulière qu'est la construction, nous devons souligner que dans cette conception, aucune distinction nette n'est faite entre lexique et syntaxe. En effet, les constructions lexicales et les constructions syntaxiques sont de même structure représentationnelle, seuls les distinguent leur complexité et leur degré d'abstraction, les constructions syntaxiques demeurant plus complexes et moins informées phonologiquement, étant donné qu'elles comprennent des *slots*.

3.2. Développement des séquences déterminant + nom

Dans le scénario développemental que nous proposons³, nous nous attachons aux étapes d'acquisition des séquences *déterminant + nom* en contexte de liaison. Pour ce faire, nous reprenons à notre compte l'organisation des constructions verbales avancée par Tomasello (2003) qui intègre la construction comme unité de base et applique les principes des modèles basés sur l'usage.

Dans une étape précoce (vers l'âge d'1 an), l'enfant récupérerait dans l'input des séquences lexicales concrètes qu'il mémoriserait dans son lexique. Elles seraient de formes diverses, plus ou moins longues, mais toutes concrètes. Des exemples de constructions de cette étape pourraient être : *l'ours*, *les ânes*, *chocolat*, etc., certaines de ces constructions pouvant contenir une CL. Lors de cette étape, l'enfant ne commettrait pas d'erreurs de liaison à proprement parler, étant donné qu'il reproduit les séquences telles qu'il les a entendues. Pour illustrer cette étape caractéristique des productions des enfants de moins de 2 ans, nous ne disposons pas de données personnelles. Cependant, Wauquier-Gravelines, qui soutient également l'hypothèse d'une étape précoce, mais en défendant d'autres interprétations⁴, fournit quelques exemples relevés chez des enfants autour de 1;6 : [əwɛzú:] pour *oiseau*, [ətɛlə] pour *pantalon* (Wauquier-Gravelines & Braud, 2005: 58). Ces deux exemples, dont seul le premier fait intervenir un nom à initiale vocalique, laissent apparaître la présence d'un déterminant devant le nom. Dans ces cas-là, nous proposons que l'enfant a récupéré ces séquences complètes qu'il reproduit en les adaptant à ces capacités articulatoires.

L'étape suivante se caractériserait par la combinaison de formes lexicales mémorisées, dans notre cas l'association de déterminants, qui joueraient le rôle de pivots, à des éléments hétérogènes du lexique. Par rapport à la première étape concrète, on se situerait ici dans une évolution vers l'abstraction par l'élaboration de schémas généraux peu spécifiques, appelés schémas pivots, du type : *un + X*, *deux + X*, *les + X*, etc. Les X correspondent à des emplacements libres sans aucune précision sur la nature morphologique des éléments pouvant les remplir. Ces schémas permettent d'ailleurs de rendre compte d'erreurs fréquemment relevées chez les enfants de 2-4 ans : les erreurs par remplacement (*un zours* /œzurs/, la forme /zurs/ intègre le schéma *un + X*) ou les erreurs par omission (*deux#ours* /dœurs/, la forme /urs/ intègre le schéma *deux + X*).

³ Le format de cet article nous oblige à présenter ici seulement les grandes lignes du scénario développemental, sans l'ensemble des données qui a permis de le construire (pour plus de précisions, voir Chevrot, Dugua & Fayol, 2005; Dugua, 2005).

⁴ Notamment, l'idée que des processus morphologiques abstraits interviennent précocement dans le développement.

La troisième étape (autour de 4-5 ans) correspondrait à la mise en œuvre de schémas plus spécifiques qui intégreraient une information abstraite de type morphologique sur la forme du Mot2. Cette information renverrait à la nature de la CL. Ces schémas seraient alors du type : *un + nX*, *deux + zX*. Généralisables à l'ensemble des formes intégrant le *slot*, ces schémas abstraits permettent ainsi de rendre compte des erreurs par surgénéralisation repérées sur des mots à consonne initiale fixe comme *un nèbre* (/œnebr/) pour *un zèbre*, *deux zombrils* (/døzõbril/) pour *deux nombrils*.

Il est essentiel de souligner que les trois étapes présentées ici sont imbriquées les unes dans les autres puisque chaque étape émerge des précédentes. De plus, les constructions caractéristiques de chacune d'elles se superposent comme des strates, les schémas de plus en plus abstraits ne supplantant pas les constructions concrètes préalables.

4. Hypothèse de travail

À partir du scénario développemental présenté ici, et selon le principe du rôle central des effets de la fréquence sur l'usage et le développement du langage, nous pouvons formuler l'hypothèse de notre présente recherche. Nous tenterons ici de mettre en lumière ces effets de fréquence sur l'acquisition des constructions *déterminant + nom* en prenant appui sur l'orientation majoritaire vers un usage au singulier ou au pluriel de différents noms.

Nous formulons notre hypothèse comme suit : plus une forme lexicale est fréquente dans la parole environnant l'enfant (l'input), plus elle sera ancrée dans son lexique (Langacker, 1987) et plus elle devrait apparaître dans les erreurs. Par exemple, si la forme /nurs/ est plus fréquente dans l'input que la forme /zurs/, on devrait recueillir plus d'erreurs de type *deux nours* (/dønurs/) que d'erreurs de type *un zours* (/œzurs/).

Par ailleurs, nous observerons si une évolution apparaît au fil du développement. Au vu des étapes développementales soulignées par notre scénario, nous pensons que le lien entre forme perçue et forme produite devrait être fort chez les plus jeunes enfants puis qu'il devrait s'affaiblir dès que les schémas spécifiques deviennent les plus productifs. Dans les étapes tardives, ce ne serait plus la fréquence des formes lexicales qui influencerait les productions mais la productivité des schémas.

5. Méthodologies

Nous avons mis en place une double méthodologie pour apporter des éléments de réponse à notre hypothèse : la première permettra d'accéder aux fréquences relatives de séquences *déterminant + nom* dans la parole adulte et la seconde de recueillir des productions enfantines sur ces mêmes séquences.

5.1. Test d'intuition auprès de locuteurs adultes

À notre connaissance, il n'existe pas de corpus oraux de parole adulte suffisamment importants pour calculer les fréquences *déterminant + nom* à voyelle initiale dont nous avons besoin. C'est pourquoi nous avons construit un test de jugement de fréquence. Soixante-dix étudiants de première année de Sciences du Langage ont pris part à cette tâche. Nous avons testé 24 Mots2 à initiale vocalique⁵ et 6 à initiale consonantique servant de distracteurs, et ce suivant deux modalités :

- Modalité « définie » : les séquences à juger se présentent sous la forme *l' + Mot2* vs *les + Mot2*, ou l'inverse *les + Mot2* vs *l' + Mot2*

⁵ Dans le cadre de cette recherche, nous utiliserons les résultats de 8 d'entre eux, ceux qui ont été testés dans des productions enfantines.

- Modalité « indéfinie » : les séquences à juger se présentent sous la forme *un* + Mot2 vs *des* + Mot2 ou l'inverse *des* + Mot2 vs *un* + Mot2

Chaque Mot2 a été jugé à deux reprises, une fois dans chaque modalité. Ainsi, nous disposons de 140 jugements par mot. L'ordre singulier/pluriel des séquences est contrebalancé avec la modalité : si un mot est présenté dans l'ordre pluriel vs singulier dans une modalité (par exemple : *les ours* vs *l'ours*), il sera présenté suivant l'ordre singulier vs pluriel dans l'autre modalité (par exemple *un ours* vs *des ours*).

La passation était collective, chaque participant disposait d'une feuille de réponses avec toutes les séquences. Les paires de séquences étaient lues par un expérimentateur et les participants devaient cocher la forme qu'ils pensaient entendre le plus souvent (en opposition à l'autre).

5.2. Expérimentation de dénomination d'images

Par l'expérimentation de dénomination d'images, nous avons cherché à recueillir un grand nombre de productions enfantines en contexte *déterminant* + *nom*. L'échantillon se compose de 511 enfants âgés de 2;3 à 6;3 et répartis en quatre tranches d'âge, comme suit.

Tranche d'âge	Échelle d'âge	Nombre d'enfants
Tranche d'âge 1	2;3-3;1	88
Tranche d'âge 2	3;2-4;1	131
Tranche d'âge 3	4;2-5;0	149
Tranche d'âge 4	5;1-6;3	143

Tableau 1 - Répartition des enfants de l'expérimentation en 4 tranches d'âge

Ce large échantillon a été obtenu grâce au regroupement de trois expérimentations basées sur un protocole similaire avec des échantillons différents (200, 189 et 122 enfants). Au niveau de la procédure, ne différait que le choix des Mots2. Dans l'expérimentation impliquant 200 enfants, 4 Mots2 étaient utilisés : *arbre*, *écureuil*, *éléphant*, *ours* (Dugua, 2006) ; dans celle avec 189 enfants, 6 Mots2 : *arbre*, *avion*, *éléphant*, *escargot*, *ordinateur*, *ours* (Nardy, 2003) ; et dans la dernière, avec 122 enfants, 8 Mots2 : les 6 Mots2 de l'expérimentation précédente ainsi que *œil* et *œuf*. Par conséquent, tous les Mots2 n'ont pas été produits par le même nombre d'enfants ; cependant, pour les analyses, nous avons globalisé les productions, sans tenir compte de cette donnée.

Deux conditions organisent cette expérimentation : l'une où le nom doit être dénommé dans le contexte *un* + Mot2 et l'autre dans le contexte *deux* + Mot2. Pour ce faire, chaque image représentant les Mots2 était dessinée en simple et double exemplaires. En outre, entre chaque image cible, était insérée l'image d'un mot distracteur à consonne initiale (par exemple, *cochon*, *camion*, *balai*, etc.).

Pour cette recherche, nous avons observé seulement les taux d'erreurs par remplacement, c'est-à-dire les erreurs dans lesquelles la forme du Mot2 n'est pas celle attendue après un Mot1 donné : par exemple, *deux ours* (/dɔ̃nurs/), *un zavion* (/œ̃zavjɔ̃/). Deux types d'erreurs seront distingués :

- (1) Les erreurs dans lesquelles la forme du Mot2 est en /n/ initial : /dɔ̃nurs/, /dɔ̃nelefã/, etc. On dira de ces erreurs qu'elles sont typiques du singulier car la consonne /n/ erronée est celle qu'on trouve après les déterminants singulier *un*, *mon*, etc.
- (2) Les erreurs dans lesquelles la forme du Mot2 est en /z/ initial : /œ̃zavjɔ̃/, /œ̃zarɔ̃r/, etc. Ces erreurs sont caractéristiques du pluriel, la consonne /z/ étant celle qui apparaît après les déterminants *les*, *deux*, *des*, *nos*, etc.

Le choix de s'appuyer sur les erreurs par remplacement s'explique en référence à deux types de schémas présentés lors du scénario développemental : les schémas généraux de type $un + X$ et les schémas abstraits tels que $un + nX$. Quand un enfant fait une erreur par remplacement, la seule possibilité qui semble envisageable est qu'il intègre une forme de Mot2 dans le *slot* d'un schéma général qui ne précise aucune information sur la nature de la forme à insérer. Dans le cadre des schémas abstraits, nous ne pouvons pas attendre d'erreurs de ce type étant donné que ces schémas portent en eux la CL attendue. En revanche, les productions justes peuvent correspondre à l'application des deux schémas. Ce n'est donc que par le biais des erreurs par remplacement que l'on a accès aux formes de Mots2 stockées dans le lexique et non à des formes déjà assimilées par les schémas.

5.3. Rapports donnant la force du pluriel

À ce premier niveau d'analyse, nous disposons du nombre des deux types d'erreurs par remplacement dans les productions enfantines pour chaque Mot2 et des scores de jugements adultes vers le pluriel ou vers le singulier pour chacun des Mots2.

Dans le but de comparer ces deux ensembles de données, nous avons mis en place un rapport donnant, pour chaque mot, la force du pluriel dans les jugements adultes et la force des erreurs en /z/ dans les dénominations enfantines.

- Dans le test d'intuition, le rapport est le suivant⁶ : [nombre de jugements en faveur du pluriel ÷ (nombre de jugements en faveur du pluriel + nombre de jugements en faveur du singulier)]. Si le rapport est proche de 0, le mot a été majoritairement jugé au singulier. Il s'agit par exemple du cas d'*arc-en-ciel* qui a un rapport à 0.036. Au contraire, le mot *habit*, jugé plus souvent pluriel, a un rapport à 0.897.
- Dans les dénominations d'images, le rapport donnant la force des erreurs en /z/ est : [nombre d'erreurs en /z/ ÷ (nombre d'erreurs en /z/ + nombre d'erreurs en /n/)].

Pour chaque Mot2, nous obtenons donc un rapport donnant sa force du pluriel dans les jugements adultes et un rapport donnant la force de l'erreur en /z/ dans les productions enfantines.

6. Résultats

Nous avons cherché si une corrélation apparaissait entre ces deux rapports, globalement, dans un premier temps, puis par tranches d'âge.

6.1. Corrélation globale

Le test de corrélation de Pearson fait apparaître une corrélation positive et significative entre les rapports de la force du pluriel dans le test d'intuition et le rapport de la force des erreurs en /z/ dans les productions enfantines ($r = .607$, $p < .0001$). Par conséquent, plus un mot est considéré par les adultes comme plus souvent utilisé au pluriel dans la parole environnante, plus l'enfant fait des erreurs impliquant la forme plurielle du mot (avec un /z/ en son initiale).

6.2. Corrélations par tranches d'âge

L'analyse par tranche d'âge permet d'observer si une évolution apparaît au fil du développement. Le tableau 2 synthétise les résultats de la corrélation de Pearson par tranche d'âge.

⁶ Pour le calcul du rapport donnant la force du pluriel dans les jugements adultes, nous avons globalisé les deux modalités « définie » et « indéfinie ».

Tranche d'âge	Échelle d'âge	Corrélation de Pearson	Significativité
Tranche d'âge 1	2;3-3;1	$r = .852, p = .015$	Significatif
Tranche d'âge 2	3;2-4;1	$r = .656, p = .055$	Tendanciel
Tranche d'âge 3	4;2-5;0	$r = .733, p = .025$	Significatif
Tranche d'âge 4	5;1-6;3	$r = .299, p = .434$	Non significatif

Tableau 2 - Résultats de la corrélation de Pearson par tranche d'âge

La corrélation est significative dans les tranches d'âge 1 et 3, tendancielle dans la tranche d'âge 2 et non significative dans la dernière.

7. Conclusion et discussion

Les résultats de cette recherche suggèrent que l'orientation singulier/pluriel de certains noms dans la parole influence les erreurs par remplacement faites par les enfants en contexte de liaison. Les données par tranches d'âge font apparaître que ces effets ne s'observent plus chez les enfants de 5-6 ans.

En termes de schémas de construction, ces résultats suggèrent que la forme insérée dans les schémas précoces, tels *un* + X et *deux* + X, est influencée par sa fréquence dans la parole environnante. Plus précisément, si un Mot2 est plus souvent employé au pluriel, la forme lexicale en /z/ initial de celui-ci sera plus fréquemment insérée après le déterminant *un*, que la forme en /n/ initial (plus typique du singulier) après le déterminant *deux*. Plus tard, lorsque les schémas se spécifient autour du déterminant, la manifestation des effets de fréquence n'est plus visible en production. En effet, les schémas *un* + nX et *deux* + zX, qui deviennent les plus productifs, portent en eux la marque « morphologique » de la consonne de liaison attendue. Ainsi, ils assimilent les formes lexicales entrantes dans les *slots* sous la forme déterminée par le schéma. Ce processus rend compte d'une part, de la disparition des erreurs par remplacement après 5 ans (Dugua, 2005) et d'autre part, de la diminution de l'influence de la fréquence en production.

Les résultats présentés dans cette recherche apportent des indices probants sur le fonctionnement et l'évolution parallèles du lexique et de la syntaxe des jeunes enfants. En effet, ils vont dans le sens de l'hypothèse que nous soutenons et selon laquelle lexique et syntaxe se situent sur un continuum ; le lexique étant caractérisé par des procédés concrets, directement influencés par l'input et la syntaxe par un traitement de plus en plus abstrait des éléments hétérogènes du lexique. L'image du continuum se traduit alors par l'émergence des schémas abstraits caractéristiques de la syntaxe à partir des schémas concrets ou des items stockés dans le lexique. Imbriqués l'un à l'autre, ces deux niveaux linguistiques suivent une même trajectoire développementale. La fréquence qui influence clairement l'acquisition du lexique se retrouve, de fait, également dans le fonctionnement précoce de la syntaxe. L'ensemble de cette recherche nous encourage à ne pas cloisonner les niveaux linguistiques dans nos analyses. Comme nous l'avons montré, un phénomène considéré majoritairement comme phonologique peut apporter un éclairage novateur sur les niveaux lexicaux et syntaxiques.

Références

- BOË, Louis-Jean & TUBACH, Jean-Pierre (1992). *De A à Zut: dictionnaire phonétique du français parlé*. Grenoble : Ellug.
- BYBEE, Joan (2003 [2001]). *Phonology and language use*. Cambridge : Cambridge University Press.
- CAMERON-FAULKNER, Thea, LIEVEN, Elena & TOMASELLO, Michael (2003). A construction based analysis of child directed speech, *Cognitive Science*, 27: 843-873.

- CHEVROT, Jean-Pierre & FAYOL, Michel (2000). L'acquisition de la liaison: enjeux théoriques, premiers résultats, perspectives, *Lidil*, 22: 11-30.
- CHEVROT, Jean-Pierre, DUGUA, Céline & FAYOL, Michel (2005). Liaison et formation des mots en français: un scénario développemental, *Langages*, 158: 38-52.
- DIESSEL, Holger (2004). *The acquisition of complex sentences*. Cambridge : Cambridge University Press.
- DUGUA, Céline (2005). De la liaison à la formation du lexique chez les jeunes enfants francophones, *Le Langage et l'Homme*, 40: 163-182.
- DUGUA, Céline (2006). Liaison, segmentation lexicale et schémas syntaxiques entre 2 et 6 ans. Un modèle développemental basé sur l'usage. Thèse de doctorat, Université Stendhal Grenoble3, Grenoble.
- GOLDBERG, Adele E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago and London: The University of Chicago Press.
- LANGACKER, Ronald W. (1987). *Foundations of cognitive grammar*. Vol.I. Stanford : Stanford University Press.
- LIEVEN, Elena, BEHRENS, Heike & TOMASELLO, Michael (2003). Early syntactic creativity: a usage-based approach, *Journal of Child Language*, 30: 333-370.
- MORIN, Yves-Charles (2003 [1998]). Remarks on prenominal liaison consonant in French, in *Living on the Edge - 28 Papers in Honour of Jonathan Kaye*, Ploch, Stephan (Eds). Berlin : Mouton de Gruyter, 385-400.
- NARDY, Aurélie (2003). Production et jugement d'acceptabilité entre 2 et 6 ans : Aspects psycholinguistiques et sociolinguistiques de l'acquisition des liaisons. Mémoire de DEA, Grenoble : Université Stendhal.
- TOMASELLO, Michael (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge [Massachusetts] : Harvard University Press.
- WAUQUIER-GRAVELINES, Sophie (2002). Segmentation de la chaîne parlée et détermination des frontières gauches dans l'acquisition du lexique précoce, *communication au GDR Phonologies*. Grenoble : Université Grenoble.
- WAUQUIER-GRAVELINES, Sophie (2003). Du réalisme des formalisations phonologiques contemporaines: que nous apprennent les données d'acquisition?, in *Phonologie: champs et perspectives*, Angoujard, Jean-Pierre & Wauquier-Gravelines, Sophie (Eds). Paris : Presses de l'Ecole Normale Supérieure de Fontenay-St Cloud, 9-34.
- WAUQUIER-GRAVELINES, Sophie & BRAUD, Virginie (2005). Proto-déterminant et acquisition de la liaison obligatoire en français, *Langages*, 158: 53-65.

MODALISATION ET APPROXIMATIONS SÉMANTIQUES ENTRE 2 ET 8 ANS : POUR UNE DICHOTOMIE NOMS VS VERBES

Juliette ÉLIE^{a et b}

elie@univ-tlse2.fr

^aLaboratoire Jacques Lordat

Pavillon de la recherche

Université Toulouse le Mirail

5 allées Antonio Machado

31058 Toulouse

^bCERPP (Centre d'Etude et de Recherche en PsychoPathologie)

Université Toulouse le Mirail

5 allées Antonio Machado

31058 Toulouse

Abstract : Children present an ability to modalise nominal semantic approximation ("the lollipop it is like an umbrella") during early acquisition (2-4 years old) but do not use this modalisation within the verbal framework before the age of 8. This would tend to show that the categorization of verbs is less "rigid" and is acquired later than that of referential names.

1. Introduction

Le propos de cet article est de montrer en quoi la modalisation d'énoncés de type métaphorique chez l'enfant permet de soutenir l'idée d'une dichotomie noms référentiels vs verbes d'action.

Pour cela, il nous sera nécessaire, dans un premier temps, de présenter les phénomènes linguistiques ici en jeu. Nous rappellerons donc, tout d'abord, en quoi consiste la métaphore et comment elle se manifeste chez l'enfant, puis, en exposant nos hypothèses, nous expliquerons ce que l'on entend par modalisation et plus précisément par modalisation de métaphore, avant de pouvoir faire les premières conclusions de notre étude en cours.

2. Problématique

2.1. La métaphore

La métaphore est une figure de style que l'on définit de manière générale comme une dénomination inhabituelle alors même qu'est disponible en langue une entité lexicale conventionnelle, ainsi que le montre l'exemple suivant :

Achille est un lion

Le sens de l'énonciation correspondante pourrait être :

Achille est fort et courageux

On peut parler ici de métaphore car on substitue un terme à un autre terme, termes qui sont reliés par une relation de ressemblance ou de conformité ; le terme qui est absent de l'énoncé est restituable sur l'axe paradigmatique :

Achille est un lion
.....
être fort et courageux

2.2. La métaphore chez l'enfant

Notre intérêt pour la métaphore nous a permis de mettre en rapport deux phénomènes, la catégorisation et la lexicalisation, phénomènes qui sont intrinsèquement liés et particulièrement cruciaux entre 2 et 4 ans (Waxman & Gelman, 1986). En effet, cette mise en rapport est primordiale dans l'investigation de la métaphore puisque cette dernière se définit plus précisément comme un franchissement de frontières catégorielles, tant en psycholinguistique (Gardner & Winner, 1979; Fourment-Aptekman, 1994, 1996) qu'en linguistique (Kleiber, 1999; Prandi, 2002).

2.2.1. Métaphores nominales et verbales

Néanmoins, on constate que les études en ce domaine portent quasi exclusivement leur attention sur le lexique des noms (Bernicot 1981; Tomasello & Merriman 1995, Bassano 2000). En effet, les métaphores observées se basent majoritairement sur un rapprochement entre deux objets ayant des ressemblances physiques ou fonctionnelles : ballon / lune, ballon / pomme, maïs / batte de base-ball, brindille tordue / tuyau d'essence... ainsi que le remarquent Gardner et Winner (1979) :

« ... les métaphores des enfants sont limitées aux comparaisons physiques et consistent simplement à attribuer de nouveaux noms à des objets physiques perceptibles. » (Gardner & Winner, 1979 : p. 86) ¹.

« La plupart des métaphores sont basées sur une ressemblance physique entre les éléments plutôt que sur un lien conceptuel, expressif, ou psychologique. » (Gardner & Winner, 1979 : p. 85) ².

Ces études montrent une importante production d'énoncés d'allure métaphorique de type nominal chez le jeune enfant : par exemple, l'enfant dit « ballon » pour désigner la lune (Gardner & Winner, 1979; Bredart & Rondal, 1982; Clark, 1993; Fourment-Aptekman, 1994, 1996; Lagarano, 1997).

Il n'en reste pas moins que la production d'énoncés d'allure métaphorique à pivot verbal est non négligeable chez l'enfant, comme le montre Duvignau (2002) en étudiant un corpus longitudinal d'un enfant et un prélèvement effectué chez 100 enfants. Elle souligne, à travers ce recueil d'énoncés, que

« la compétence analogique des jeunes enfants n'est pas restreinte à la seule mise en relation de propriétés d'objets physiques. Elle ne se manifeste pas seulement à travers des noms d'objet mais concerne également des actions et se manifeste à travers des énoncés mettant en jeu des verbes ou des adjectifs verbaux qui sont produits en quantité non-négligeable... » (Duvignau, 2002 : p. 226).

Ainsi on trouve un grand nombre d'exemples de production d'approximations sémantiques de type verbal chez l'enfant : par exemple, il dit « je déshabille la banane » pour dénommer l'action [épilucher une banane] (Duvignau, 2002).

2.2.2. Sur-extensions, métaphores et approximations sémantiques

Ces énoncés de type métaphorique sont le plus souvent considérés soit comme des sur-extensions erronées soit comme des métaphores dans la littérature (Lagarano, 1997). Les sur-

1 Texte original : "... the children's metaphors are limited to physical comparisons and consist of simply ascribing new names to discrete physical objects.", p. 86.

2 Texte original : "Almost all early metaphors are based on physical resemblance between elements, rather than on a conceptual, expressive, or psychological link.", p. 85.

extensions sont perçues comme des erreurs puisque l'on considère que les enfants manquent de vocabulaire et utilisent un terme inapproprié en l'absence du terme approprié :

« ... on a observé que durant les premières étapes du développement lexical, les mots étaient souvent employés par les jeunes enfants de manière non-conforme à l'usage adulte, reflétant diverses sortes d'erreurs d'extension. Au rang des plus fameuses se trouvent les erreurs de sous-extensions et de sur-extensions. » (Bassano, 2000 : p. 152-153).

Ainsi, si un enfant dit *ballon* pour *lune* ou *déshabille* pour *épluche* alors qu'il ne dispose pas des termes adéquats, il produira une sur-extension erronée. À l'inverse, s'il possède les termes appropriés et si les mots en jeu sont rangés dans des catégories conceptuelles distinctes, alors il produira des métaphores. Le jeune enfant peut donc constituer deux types d'énoncés : soit des sur-extensions, soit des métaphores.

Sur ce point, il est important de remarquer que si les lexicalisations « métaphoriques » des jeunes enfants peuvent parfois prétendre au statut de métaphore dans le cadre nominal, il n'en va pas de même pour le cadre verbal. On constate, en effet, que la dichotomie entre les noms et les verbes montre une différence intervenant essentiellement dans les relations entre les termes en jeu et donc dans le mode de structuration du lexique.

Concernant les noms référentiels, l'enfant structure son lexique par la mise en relation d'objets ayant une ressemblance physique ou fonctionnelle, comme nous avons pu le voir plus haut et on se rend compte également que les deux noms correspondants aux objets ne sont pas synonymes : par exemple, *lune* n'est pas synonyme de *ballon*, la relation mise en jeu dans une métaphore nominale est de l'ordre de la polysémie. Ainsi, certaines « lexicalisations » métaphoriques peuvent avoir le statut de métaphore.

Quant aux verbes, l'enfant structure son lexique par proximité sémantique ou co-hyponymie inter-domaine et on remarque que ces verbes en jeu sont synonymes inter-domaine. Ainsi nous avons affaire, non plus à des erreurs ou des métaphores, mais à des approximations sémantiques fondées sur une relation de synonymie inter-verbes inter-domaines. On distinguera alors les approximations sémantiques intra-domaine où le caractère approximatif de l'énoncé est lié à une *non-correspondance avec l'action dénommée* (dire « couper l'orange » pour une orange que l'on épluche) et des approximations sémantiques extra-domaine où le caractère approximatif de l'énoncé se fonde sur une *tension sémantique au niveau des termes de la langue* (dire « déshabiller l'orange » pour une orange que l'on épluche) (Duvignau, 2002). On peut se représenter ces notions d'intra et d'extra-domaine sous la forme d'un arbre comme représenté ci-dessous dans la figure 1.

La production de ce type d'énoncés chez le jeune enfant fait apparaître une compétence cognitive fondamentale présente dès le plus jeune âge : l'enfant est doté d'une capacité analogique qui lui permet de « voir la ressemblance dans la différence » (Sander, 2000). Les approximations sémantiques sous-tendent ainsi une flexibilité tant sémantique que cognitive importante chez le jeune enfant (Duvignau, 2002 ; Duvignau & Gaume, 2003).

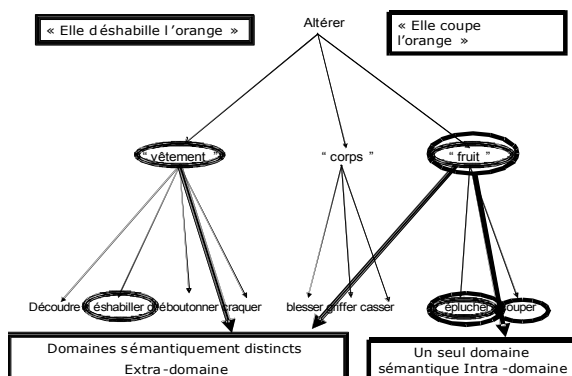


Figure 1 - Représentation schématique des notions d'intra et d'extra domaine

3. Hypothèses

Dans ce contexte, à travers l'étude de la modalisation des approximations sémantiques verbales vs nominales, nous souhaitons défendre l'idée d'une dichotomie noms (circonscrits aux noms référentiels) vs verbes d'actions, avancée par Gentner & Boroditsky (2001). Notre hypothèse est que la modalisation des approximations verbales est plus tardive que celle des approximations nominales pour laquelle la littérature établit une production dès l'âge de 3-4 ans (« La sucette c'est comme un parapluie » Joane 3;6 ans). La validation de cette hypothèse s'inscrit dans la lignée de Duvignau (2002) et permettrait d'avancer l'idée d'une catégorisation du lexique verbal plus tardive que celle du lexique nominal.

3.1. La modalisation

On définit généralement la modalisation comme une activité métalinguistique qui désigne une capacité à « nous élever au-dessus de la langue, de nous en abstraire, de la contempler » (Benveniste, 1974). Gombert (1990) décrit l'aptitude métalinguistique comme « une attitude réflexive sur les objets langagiers et leur manipulation ». Vion (2001), dans son article *Modalités, modalisations et activités langagières*, fait un inventaire des auteurs ayant traité ce sujet et choisit de définir la modalisation « comme une double énonciation avec production d'un commentaire réflexif de l'une sur l'autre » (Vion, 2001 : p 225). Il relève ainsi différents types d'énoncés relevant de la modalisation :

- « *Pierre viendra **certainement** jeudi* » vu comme un commentaire sur le « dit », il exprime un jugement, une attitude modale.
- « *Il a atterri, **si je puis dire**, à l'université* » considéré comme un commentaire sur le « dire » et plus précisément comme une glose méta-énonciative qui porte sur le choix des mots et les manières de s'exprimer.
- « *Pourrais-tu, **sans vouloir te commander**, aller voir si le facteur est passé ?* » lui aussi vu comme un commentaire sur le « dire », à la différence qu'il s'agit non plus d'une glose méta-énonciative mais d'une glose métadiscursive c'est-à-dire que ce type de commentaire porte sur la valeur du dire et la façon de l'interpréter.

3.2. La modalisation de la métaphore

On peut constater la présence assez régulière de termes tels que « vrai », « un genre de », « tout », « littéralement », etc. accompagnant la métaphore. Tamba (1994) désigne ces termes comme « des modalisateurs graduels de l'énonciation figurative ». Elle distingue les modalisateurs caractérisant la métaphore et ceux caractérisant la métonymie (« tout », « entier »...). On peut en repérer deux types dans les métaphores :

- « vrai », « véritable », « même », « genre », « espèce », « sorte de »,... marquant une fonction de « notion type ».
- « littéralement », « véritablement », « vraiment », « pour ainsi dire », marquant une fonction de « prédicatif ».

« Comme » est un exemple de modalisateur qui, dans une métaphore nominale, aura pour fonction de mettre en ressemblance les éléments présents, alors que dans la métaphore verbale, il aura un rôle atténuateur (ex : « je fus comme inondé de lettres... »). Les marques de modalisation présentes dans les métaphores servent la plupart du temps à atténuer l'in vraisemblance de la métaphore ou à marquer, au contraire, son caractère abusif.

3.3. La modalisation de la métaphore chez l'enfant

Le jeune enfant est-il capable de modaliser un énoncé d'allure métaphorique ? Tamine (1988) précise que « le développement de l'activité métalinguistique est lié à l'émergence et aux progrès de l'activité métalinguistique et à la prise de conscience du signe linguistique » ; en effet, produire une métaphore sous-entend qu'il faut avoir une intention de changer de sens en jouant sur des traits sémantiques. Ainsi, la modalisation d'une métaphore montrera que l'enfant a conscience qu'il utilise un terme différent de celui qu'on pourrait attendre pour dénommer un objet ou désigner une action. Voyons alors comment vérifier si l'enfant modalise ces énoncés de type métaphorique.

4. Méthodologie

Notre corpus est constitué de 100 approximations sémantiques nominales et verbales issues du discours enfantin recueillies, d'une part, à partir d'un corpus longitudinal d'un enfant et d'un prélèvement effectué chez 100 enfants (Duvignau, 2002), et d'autre part, via le site internet *Enfandises*, site dédié aux mots d'enfants (Fourel, 2004, 2005). Notre corpus se compose également de 200 approximations sémantiques verbales recueillies en situation expérimentale auprès de 80 enfants âgés de 2 à 8 ans. Ce protocole expérimental, qui se déroule à l'oral, consiste en une tâche de dénomination de 17 vidéos d'actions « qu'est ce qu'elle a fait la dame ? » suivie d'une tâche de reformulation : « ce qu'elle a fait la dame, dis-le moi d'une autre manière, avec d'autres mots ».

5. Premiers résultats

Les premières analyses effectuées montrent que l'enfant présente une capacité de modalisation durant l'acquisition précoce (2-4 ans) qui lui permet de marquer le caractère inhabituel des approximations sémantiques nominales :

- « Maman tu es aussi jolie qu'un quartier de lune » (Victoire : 4 ans).
- Guillaume est très intrigué par les nombreux grains de beauté présent sur la peau de sa grand-mère : « oh, mais il y en a partout s'écrit-il, tu es une vraie girafe » (Guillaume : 3 ans).
- Julia regarde une coccinelle : « on dirait qu'elle a attrapé la varicelle ! » (Julia : 4 ans).
- « La sucette c'est comme un parapluie » (Joane : 3 ans et 6 mois).
- Voyant sa petite sœur avec 2 couettes : « Euh ! et ben, Candice, on dirait un cerf !!! » (Laurène : 3 ans).

En revanche, malgré une forte production d'approximations sémantiques de type verbal (« elle déshabille la banane » par exemple), aucune d'entre elles ne portent de marque de modalisation. Ce résultat permet de soutenir l'idée que le jeune enfant n'établit pas de contraste fort entre des verbes qui véhiculent un même noyau de sens (« casser » et « déchirer » qui renvoient tous deux à la notion /Détériorer/) : ces verbes détiendraient la même valeur intensionnelle et communicationnelle, ils seraient mis en équivalence, ce qui

expliquerait que l'enfant puisse les utiliser de manière quasi-indifférente. Ainsi, durant la période du développement lexical précoce, l'approximation sémantique « il a cassé le livre » pour l'action [déchirer un livre] manifeste un rapprochement entre : « casser » et « déchirer », verbes qui seraient moins contrastés du point de vue de leur sens que les items nominaux renvoyant à des référents concrets comme « champignon » et « parapluie ». L'établissement de frontières catégorielles strictes entre des verbes proches sémantiquement ressort donc comme plus tardive chez le jeune enfant (2-4 ans). Ce n'est que vers 8 ans que nos corpus permettent d'observer l'apparition de premières modalisations d'approximations sémantiques verbales du type : « **on dirait qu'elle** a coupé le papier » (Duvignau & Elie, en préparation).

6. Conclusion

Le jeune enfant a donc une capacité analogique permettant de produire des approximations sémantiques tant nominales que verbales qui jouent un rôle fondamental dans la constitution progressive du lexique chez le jeune enfant et qui sont des indices de sa flexibilité cognitive.

Cependant, le fait que l'enfant ne manifeste pas de modalisation d'énoncé d'allure métaphorique à pivot verbal avant l'âge de 8 ans alors qu'il présente une modalisation d'approximations sémantiques à foyer nominal dès l'âge de 3 ans, tend à montrer que la catégorisation des verbes est moins « rigide », moins « stricte » que celle des noms référentiels et donc plus tardive (Duvignau, 2002). Il nous reste à rappeler que cette étude concernant la modalisation de la métaphore chez le jeune enfant n'en est qu'à ces débuts et qu'il nous faudra encore vérifier plusieurs points, entre autres l'influence de la demande de reformulation – l'enfant modalise-t-il son énoncé parce qu'on lui a demandé ou parce qu'il ne peut faire autrement ? Ces premières investigations soutiennent, néanmoins, notre hypothèse qu'il faudra ainsi étayer et compléter, notamment en mettant en place une étude longitudinale de la production d'approximations sémantiques entre 3 et 8 ans.

Références

Références bibliographiques

- BASSANO, Dominique (2000). La constitution du lexique : Le « développement lexical précoce », in *L'acquisition du langage, Le langage en émergence. De la naissance à trois ans*, Fayol, Michel & Kail, Michèle. Paris : PUF, 137-169.
- BENVENISTE, Emile (1974). *Problèmes de linguistique générale*. Paris : Gallimard.
- BERNICOT, Josie (1981). *Le développement des systèmes sémantiques de verbes d'action*. Paris : Monographies Françaises de Psychologie, CNRS Editions.
- BREDART, Serge, RONDAL, Jean-Adolphe (1982). *L'analyse du langage chez l'enfant : les activités métalinguistiques*. Bruxelles : Mardaga.
- CLARK, Eve (1993). *The lexicon in acquisition*. Cambridge : Cambridge University Press.
- DUVIGNAU, Karine & GAUME, Bruno (2003). Métaphore verbale et approximation, in *Regards croisés sur l'analogie*, Duvignau, Karine, Gasquet, Olivier & Gaume, Bruno (Eds). RIA, vol n°5/6, Paris : Hermès Lavoisier, 701-704.
- DUVIGNAU, Karine (2002). *La métaphore, berceau et enfant de la langue. La métaphore verbale comme approximation sémantique par analogie dans les textes scientifiques et les productions enfantines (2-4 ans)*, Thèse de Sciences du Langage, Université Toulouse Le-Mirail.
- DUVIGNAU, Karine & ELIE, Juliette (en préparation). Capacité de modalisation chez le jeune enfant : vers une dichotomie noms versus verbes.
- FOUREL, Sandrine (2004). *Dichotomie noms versus verbes chez le jeune enfant via la modalisation de la métaphore*, Mémoire de maîtrise de Sciences du Langage, Université Toulouse Le Mirail.

- FOUREL, Sandrine (2005). *Approximations sémantiques nominales vs verbales*, Mémoire de DEA de Sciences du Langage, Université Toulouse Le Mirail.
- FOURMENT-APTEKMAN, Marie-Claude (1994). *La métaphore chez l'enfant : étude des relations entre production et compréhension au cours du développement*, Thèse d'Etat, Université Paris V.
- FOURMENT-APTEKMAN, Marie-Claude (1996). La compréhension des métaphores et des pseudo-métaphores chez des enfants âgés de 4 à 8 ans, *L'Année Psychologique*, n°3 : 443-459.
- GARDES-TAMINE, Joëlle. (1988). Les conditions d'acquisition de la métaphore chez l'enfant, *Recherches sur la Philosophie et le Langage*, n°9 : 260-271.
- GARDNER, Howard & WINNER, Ellen (1979). The child is father to the metaphor, *Psychology Today*, May : 81-91.
- GENTNER, Drede & BORODITSKY, Lera (2001). Language acquisition and conceptual development, in *Cognition and Categorization*, Bowerman, Melissa & Levinson, Stephen (Eds). Cambridge : Cambridge University Press, 215-256.
- GOMBERT, Jean-Emile (1990). *Le développement métalinguistique*. Paris : Presses universitaires de France.
- KLEIBER, Georges (1999). Une métaphore qui ronronne n'est pas toujours un chat heureux, in *La métaphore entre philosophie et rhétorique*, Charbonnel, Nanine & Kleiber, Georges. Paris : PUF : 83-135.
- LAGARANO, Marina (1997). Production et compréhension des métaphores chez l'enfant : de la similitude à la métaphore, *Archives de Psychologie*, 253 : 141-165.
- PRANDI, Michèle (2002). La métaphore : de la définition à la typologie, *Langue française*, 134: 6-21.
- SANDER, Emmanuel (2000). *L'analogie, du Naïf au Créatif. Analogie et catégorisation*. Paris : L'Harmattan.
- TAMBA, Irène (1994). Une clé pour différencier deux types d'interprétation : figure, métaphorique et métonymique, *Langue Française*, 101 : 26-34.
- TOMASELLO, Mickael & MERRIMAN, William (Eds) (1995). *Beyond Names for Things: Young Children's Acquisition of Verbs*. Hillsdale : Lawrence Erlbaum Publishers.
- VION, Robert (2001). Modalités, modalisations et actives langagières, *Marges Linguistiques*, 2: 209-231.
- WAXMAN, Sandra R. & GELMAN, Rochel (1986). Preschoolers' use of superordinate relations in classification and language, *Cognitive Development*, 1: 139-156.

Références aux sites Internet

DARJEELINK, Web strategies & social software
<http://www.enfandises.com>
Consulté en Juin 2006.

LE DÉVELOPPEMENT MÉTAPHONOLOGIQUE CHEZ LES ENFANTS BILINGUES ET MONOLINGUES

Angélique LAURENT

angelique.laurent@univ-tlse2.fr

Université Toulouse Le Mirail

Laboratoire Cognition, Communication et Développement

Pavillon de la Recherche

5, allées Antonio Machado

31058 Toulouse Cedex 9

Abstract : Many authors have suggested that bilingualism influences the development of phonological skills. In a monolingual context, these abilities develop in interaction with reading. This research attempts to compare the development of phonological skills in bilingual and monolingual before reading. In other words, does bilingualism improve phonological awareness?

1. Cadre théorique

Depuis quelques années, l'étude du bilinguisme connaît un essor important. Une des premières raisons à cela est la prise de conscience que la population bilingue représente une grande partie de la population mondiale : en témoigne le nombre de langues recensées dans le monde, entre 3000 et 7000. Ainsi, pour décrire et expliquer le développement cognitif et langagier, il convient de prendre en compte les performances des bilingues autant que celles des monolingues.

Mais la difficulté la plus importante, quand on s'attache à l'étude de cette thématique, est de retenir une définition du bilinguisme compte tenu de la quantité et de la diversité des définitions présentes dans la littérature. Les auteurs proposent des définitions selon leur champ de recherche, mettant en avant des caractéristiques linguistiques, psychologiques, culturelles, sociologiques, etc. des personnes bilingues. De manière générale, on peut représenter cette diversité définitoire sous la forme d'un continuum composé, à une extrémité, d'une définition la plus large possible – MacNamara (1967) : une personne bilingue est une personne qui possède une compétence minimale dans une des quatre habiletés linguistiques (comprendre, parler, lire et écrire) dans une langue autre que sa langue maternelle – à l'autre, une définition la plus stricte possible – Bloomfield (1933) : le bilinguisme est défini comme la possession d'une compétence de locuteur natif dans les deux langues. Une multitude de définitions peut être positionnée sur ce continuum selon le degré de compétences dans chaque langue.

Dans notre perspective, plutôt que de reprendre une définition selon des caractéristiques psychologiques, culturelles, sociologiques, etc., il nous semble plus judicieux de retenir la distinction soutenue par certains auteurs (Deshays, 1998) qui définissent le bilinguisme par rapport à la manière dont les deux langues sont apprises par l'enfant :

- le bilinguisme simultané : c'est le cas des enfants qui sont en contact avec deux langues dès la naissance et les apprennent simultanément sans difficulté, par exemple, lorsque les parents sont de langues maternelles différentes et utilisent leur langue respective avec l'enfant.
- le bilinguisme consécutif : c'est le cas des enfants qui grandissent dans une famille monolingue et découvrent une deuxième langue très rapidement dans la petite enfance,

par exemple au début de la scolarisation, lorsque la langue de l'école est différente de la langue de la maison.

Parmi toutes les recherches sur le développement langagier, nombreuses sont celles qui témoignent de différences plus ou moins marquées entre les enfants bilingues (simultanés ou consécutifs) et les enfants monolingues sur le plan syntaxique, sémantique, lexical et phonologique (Bialystok, 1991; Perregaux, 1994; Mann & Wimmer, 2002, etc.). Ces différentes études montrent un certain avantage des enfants bilingues par rapport aux enfants monolingues mais qui est fonction, d'une part, du niveau de maîtrise des langues et, d'autre part, du système d'écriture (phonographique ou logographique). Ces études examinent plus particulièrement les compétences linguistiques des enfants lors de l'apprentissage de la lecture et de l'écriture. Plus rares sont les études qui s'intéressent aux bilinguismes des enfants prélecteurs, donc avant l'apprentissage de la lecture. Bruck & Genesee (1995) ou Campbell & Sais (1995), par exemple, montrent que des enfants bilingues consécutifs (anglais/français et anglais/italien respectivement) âgés de 4-5 ans ont des performances en analyse segmentale de la parole supérieures aux enfants monolingues. Les auteurs font état d'un avantage métalinguistique non négligeable à des tâches de conscience phonologique en faveur des enfants bilingues.

Or, il ressort d'autres recherches (Gombert, 1990, Gombert & Colé, 2000) que dans ce type de compétences métalinguistiques, les capacités métaphonologiques (capacité à prendre en compte, à identifier les unités phonologiques et à les manipuler de façon délibérée, Gombert, 1990) se développent en interaction avec l'apprentissage de la lecture. En effet, les capacités métaphonologiques, mesurées au moyen de tâches de suppression de phonème par exemple, se développent dans le contexte de la lecture ou bien dans le contexte d'entraînements spécifiques. Les enfants prélecteurs ont d'ordinaire des performances très faibles quand il s'agit de supprimer le phonème initial d'items monosyllabiques comparativement aux enfants qui apprennent à lire et à écrire (Content, Kolinsky, Morais & Bertelson, 1986; Martinot & Gombert, 1996).

2. Problématique

Sur la base des études précitées, il semble intéressant de s'interroger sur la nature des capacités phonologiques des enfants possédant un bilinguisme précoce simultané. En effet, si les enfants bilingues consécutifs font preuve de capacités métaphonologiques avant d'apprendre à lire et à écrire (Bruck & Genesee, 1995; Campbell & Sais, 1995), il est possible que les enfants bilingues précoces simultanés soient aussi dotés de ces capacités avant l'apprentissage de la lecture.

Dès 1934, Vygotsky suggère que l'utilisation de deux langues peut faciliter l'accès à la conscience linguistique. Il affirme que l'apprentissage d'une deuxième langue peut jouer un rôle dans l'élaboration d'une distanciation précoce entre le signifiant et le signifié. Selon lui, l'enfant prend plus rapidement conscience des formes linguistiques, de l'existence de deux formes (au moins) pour exprimer un même concept et utilise plus consciemment le mot en tant qu'instrument de pensée et expression de concept. On retrouve cette idée chez plusieurs auteurs, notamment Bialystok (1991) qui reprend le terme de « conscience de l'arbitraire ».

Partant de ces différents arguments, nous avons mené une recherche auprès d'enfants bilingues français/anglais pour tester les hypothèses suivantes. Tout d'abord, les enfants bilingues ont un développement métaphonologique plus précoce que leurs pairs monolingues, c'est-à-dire un développement métaphonologique en dehors du contexte de l'apprentissage de la lecture. Ainsi, nous postulons que les enfants bilingues feront preuve de capacités métaphonologiques avant d'apprendre à lire. De plus, du fait de l'interaction entre les deux langues acquises, les enfants bilingues ne témoigneront pas du même type de préférence pour les unités subsyllabiques que les enfants monolingues. Nous supposons

qu'ils ne feront pas preuve de la même sensibilité phonologique que leurs pairs monolingues. L'objectif est donc de déterminer la nature des capacités phonologiques des enfants bilingues, comparativement à des enfants monolingues. En d'autres termes, l'acquisition précoce d'une deuxième langue nous semble susceptible d'avoir une incidence sur le développement phonologique.

3. Méthodologie

3.1. Population

L'échantillon est constitué de 50 enfants, 30 monolingues français (15 filles et 15 garçons) et 20 enfants bilingues français/anglais (9 filles et 11 garçons). Nous entendons par enfant monolingue un enfant qui utilise et entend uniquement la langue française, dans le contexte familial et scolaire. Au contraire, l'enfant bilingue est en contact avec deux langues : le français et l'anglais, et, en référence à la définition du bilinguisme précoce simultané, utilise les deux langues à la maison, "le principe 'une personne - une langue'" étant respecté. Suite à des entretiens auprès des parents des enfants bilingues, il apparaît que ces derniers utilisent de manière équivalente les deux langues à la maison. Tous les enfants sont scolarisés en classes de grande section de maternelle, dans des écoles traditionnelles de la région toulousaine. La classe de grande section a été choisie car elle précède la classe de cours préparatoire, moment où les capacités métaphonologiques sont nécessaires pour appréhender l'apprentissage de la lecture. Par conséquent, les enfants sont pré-lecteurs, c'est-à-dire qu'ils n'ont pas encore commencé l'apprentissage formel de la lecture et de l'écriture. L'âge moyen du groupe monolingue est de 5 ans et 10 mois (de 5 ans et 5 mois à 6 ans et 4 mois) et l'âge moyen du groupe bilingue est 5 ans et 11 mois (de 5 ans et 6 mois à 6 ans et 4 mois). La plupart des parents des enfants bilingues travaillant dans l'industrie aéronautique de Toulouse, le niveau socio-économique des parents des enfants monolingues a été apparié à celui des parents des enfants bilingues. Enfin, des entretiens avec les enseignants ont permis de vérifier qu'aucun enfant n'avait de retard scolaire ou langagier particulier.

3.2. Matériel

Les enfants des deux groupes ont été testés au moyen de deux tâches. Tout d'abord, une tâche de segmentation contrainte a été proposée. Il s'agit de supprimer le phonème initial d'items monosyllabiques et de restituer oralement le segment restant ("Si je te dis *car*, tu me dis *ar*"). Cette tâche comporte 40 items du type consonne-voyelle-consonne dont 20 mots et 20 pseudo-mots, c'est-à-dire des mots inventés par l'expérimentateur. Cette tâche a été initialement utilisée par Rosner & Simon (1971) et est très souvent reprise afin de mesurer le niveau de développement des capacités métaphonologiques. Après avoir recensé les tâches métaphonologiques utilisées dans la littérature, Yopp (1988) classe la tâche de segmentation contrainte comme étant la plus difficile. L'ordre de présentation des items est aléatoire, les mots sont mélangés aux pseudo-mots. Dans un deuxième temps, les enfants ont été testés au moyen d'une tâche de préférence, tirée des travaux de Martinot & Gombert (1998). Cette tâche évalue la sensibilité des enfants aux segments subsyllabiques (attaque, début, voyelle, rime et coda) et ne nécessite pas le même niveau d'exigence cognitive que la précédente. Elle consiste à choisir, entre deux items proposés oralement, celui qui est le plus proche phonologiquement de l'item cible énoncé en premier ("*Qu'est-ce qui ressemble plus à mine? Est-ce fine ou mille?*"). Il s'agit donc d'identifier une ressemblance acoustique entre plusieurs stimuli. Il y a 40 triplets d'items dont 20 triplets de mots et 20 triplets de pseudo-mots, chaque triplet étant composé de 3 items monosyllabiques. Toutes les possibilités d'appariement ont été manipulées (cf. tableau 1). L'ordre de présentation des triplets d'items est aléatoire, les triplets de mots sont mélangés aux triplets de pseudo-mots.

Enfin, l'ordre de présentation des deux tâches est aléatoire.

Segments communs	Item cible	Item 1	Item 2
Rime - Début	Mine	Fine	<u>M</u> ille
Rime - Attaque	Mine	Fine	<u>M</u> are
Rime - Voyelle	Mine	Fine	<u>R</u> ite
Rime - Coda	Mine	Fine	<u>L</u> une
Début - Attaque	Mine	<u>M</u> ille	<u>M</u> are
Début - Voyelle	Mine	<u>M</u> ille	<u>R</u> ite
Début - Coda	Mine	<u>M</u> ille	<u>L</u> une
Attaque - Voyelle	Mine	<u>M</u> are	<u>R</u> ite
Attaque - Coda	Mine	<u>M</u> are	<u>L</u> une
Voyelle - Coda	Mine	<u>R</u> ite	<u>L</u> une

Tableau 1 - Possibilités d'appariement des segments subsyllabiques

3.3. Procédure

Chaque participant a été rencontré individuellement lors d’une session de 20 minutes environ. Les passations se sont déroulées dans l’enceinte de l’école, dans une salle proche de la salle de classe afin que les enfants puissent réaliser les tâches dans le calme.

4. Résultats

4.1. Tâche de segmentation contrainte

Les résultats ont été soumis à une analyse de variance selon le plan 2 (Langue : monolingue vs bilingue) X 2 (Lexicalité : mots vs pseudo-mots) avec des mesures répétées sur le dernier facteur et les sujets comme facteur aléatoire.

Groupes	Réponses correctes
Monolingue	5.20 (10.78)
Bilingue	6.65 (11.88)

Tableau 2 - Performances moyennes sur 40 (et écart-types) pour chaque groupe d'enfants.

Au regard du tableau 2 et comme il ressort de la littérature (Rosner & Simon, 1971; Content *et al.*, 1986; Martinot & Gombert, 1996), les enfants ont très peu réussi cette tâche, avec cependant une variation interindividuelle importante pour les deux groupes (cf. écart-types ci-dessus) qui mérite d’être soulignée. La comparaison des performances moyennes ne permet pas de mettre en évidence de différence entre les groupes, même si les enfants bilingues obtiennent une moyenne de réponses correctes très légèrement supérieure à celle des enfants monolingues. Aucun effet de la lexicalité n’a été observé, donc, quel que soit le groupe, il n’y a pas de différence entre les mots et les pseudo-mots. En d'autres termes, les enfants bilingues ne se distinguent pas des enfants monolingues lors de la résolution de cette tâche.

Cependant, l’examen des réponses erronées a permis de mettre en évidence une différence entre les deux groupes. En effet, là où les enfants monolingues restituent généralement le début des items, les enfants bilingues restituent le coda. Ainsi, ils extraient de façon non négligeable le début des items alors que les enfants monolingues extraient le coda. Par exemple, lorsque le mot "*bol*" est proposé, les enfants monolingues répondent "*bo*", et les enfants bilingues répondent "*l*". L’effet est significatif ($F(1,48) = 22.685; p = .000$).

Groupes	Réponses "Coda"
Monolingue	1.13 (3.27)
Bilingue	15.35 (15.93)

Tableau 3 - Performances moyennes sur 40 (et écart-types) des réponses "Coda" (–C) pour chaque groupe d'enfants.

Il semble donc que les enfants bilingues n’aient pas de facilité particulière pour résoudre cette tâche, néanmoins ils font preuve d’une segmentation des monosyllabes bien spécifique.

Cette segmentation particulière ne se retrouve pas chez les enfants monolingues français et n'a pas été mise en évidence, à notre connaissance, dans les études chez les enfants monolingues anglais.

4.2. Tâche de préférence

Les résultats ont été soumis à une analyse de variance selon le plan 2 (Langue : monolingue vs bilingue) X 2 (Choix : item 1 vs item 2) avec des mesures répétées pour chaque appariement. Ce traitement statistique nous a permis d'étudier la sensibilité des enfants aux différentes unités subsyllabiques. Les deux tableaux concernant la tâche de préférence doivent être lus de la façon suivante : chaque pourcentage correspond au taux du choix du segment de la ligne sur le segment de la colonne. Par exemple, la rime est "préférée" au début à 57.5% par les enfants monolingues et à 48.75% par les enfants bilingues. Apparaissent également les probabilités associées après analyse statistique.

Possibilités d'appariement	Début	Attaque	Voyelle	Coda
Rime	57.5**	66.67***	64.25**	70.83***
Début		79.18***	60**	70.82***
Attaque			37.5***	61.68***
Coda				58.32*

(Probabilité associée à l'effet : *** : $p < .0001$; ** : $p < .001$; * : $p < .01$; ns : non significatif).

Tableau 4 - Pourcentages des segments "préférés" pour chaque possibilité d'appariement pour le groupe des enfants monolingues.

Comme il ressort du tableau 4, la rime est le segment phonologique le plus saillant comparativement au début, à l'attaque, à la voyelle et au coda pour les enfants monolingues. Ceci signifie que lorsqu'on propose à un enfant monolingue de choisir entre deux segments partageant, avec l'item cible, soit la rime soit le début, l'enfant monolingue répond préférentiellement l'item partageant la rime avec l'item cible. Par exemple, le choix de l'enfant monolingue se porte sur *fine* lorsqu'il faut le comparer à *mille* et trouver celui qui ressemble le plus à *mine*. Les enfants monolingues sont sensibles, dans une moindre mesure, au début. Par conséquent, résumer les données obtenues reviendrait à établir, de façon extrapolée et sans fondement statistique, une hiérarchie des préférences dans laquelle la rime occuperait la position du segment le plus saillant, suivie du début, de la voyelle, de l'attaque et du coda.

Possibilités d'appariement	Début	Attaque	Voyelle	Coda
Rime	48.75 ns	83.75***	81.25***	90***
Début		91.25***	78.75***	83.75***
Attaque			50 ns	68.75***
Coda				73.75***

(Probabilité associée à l'effet : *** : $p < .0001$; ** : $p < .001$; * : $p < .01$; ns : non significatif).

Tableau 5 - Pourcentages des segments "préférés" pour chaque possibilité d'appariement pour le groupe des enfants bilingues.

Au regard du tableau 5, on constate que, chez les enfants bilingues, le début est légèrement plus saillant lorsqu'il est comparé aux autres segments phonologiques un à un. Lorsqu'il s'agit de sélectionner entre deux items partageant le début et la rime avec l'item cible, ces derniers choisissent plutôt l'item partageant le début avec l'item cible. Par exemple, entre les deux items *fine* et *mille*, l'enfant bilingue choisit *mille* comme item ressemblant le plus à *mine*. Après une préférence marquée pour le début, les enfants bilingues sont sensibles, dans une moindre mesure, à la rime. On pourrait donc résumer ces préférences sous une forme hiérarchique, dans laquelle le début serait le segment phonologique le plus saillant, suivi de la rime, puis de l'attaque et de la voyelle choisies dans une même proportion et enfin du coda.

De manière générale, il semble donc que tous les enfants sont plus sensibles aux grandes unités phonologiques (rime ou début) qu'aux petites unités (coda, attaque ou voyelle).

5. Discussion - Conclusion

L'objectif de cette recherche réside dans l'analyse des effets du bilinguisme sur le développement des capacités phonologiques des enfants pré-lecteurs. Les capacités phonologiques ont été étudiées au moyen d'une tâche de segmentation contrainte et d'une tâche de préférence. Les résultats obtenus vont partiellement dans le sens de nos hypothèses. La première hypothèse concernait l'avantage bilingue lors de la résolution de la tâche de segmentation contrainte. Cette tâche, qui nécessite de réelles capacités métaphonologiques pour être résolue, n'a pas été réussie par les enfants indépendamment de leur appartenance au groupe bilingue. Par conséquent, le bilinguisme ne permet pas d'avancer l'idée d'un avantage ou d'un désavantage métaphonologique. En revanche, un facteur intéressant est ressorti lors du traitement des résultats : le pourcentage d'extraction du coda, groupe consonantique final, par les enfants bilingues. Plus d'un tiers des enfants de ce groupe segmentent les items en deux parties : début/coda, alors que les enfants monolingues n'effectuent pas ce type de segmentation. Certes, le coda est une réponse erronée mais qui est plus proche de la réponses correcte, car il en fait partie intégrante, que le début, réponse donnée par les enfants monolingues. Ce résultat peut être expliqué par deux facteurs. Le premier facteur explicatif concerne la consigne donnée aux enfants, consigne constituée d'une brève explication contenant une précision sur l'objectif de la tâche et d'un exemple. L'expérimentateur indiquait aux enfants qu'il fallait restituer les sons de la fin des mots. Il est possible que les enfants bilingues se soient appuyés sur cette remarque pour personnaliser la tâche. Même si l'objectif n'a pas été atteint, ils ont essayé de résoudre la tâche d'une certaine manière, ce que n'ont pas fait les enfants monolingues. On peut se demander si un tel comportement ne reflèterait pas leurs capacités métaphonologiques. Au regard de la définition du concept, les enfants bilingues semblent opérer un travail intentionnel et réfléchi sur les items en les segmentant en deux unités, le début et le coda, et en restituant le deuxième segment à l'expérimentateur. Les enfants monolingues, pour leur part, opèrent un travail davantage automatique en donnant le début comme réponse. Extraire la partie finale d'un item et restituer la partie initiale est plus aisé qu'extraire la partie initiale et restituer la partie finale. Le second facteur explicatif de ce résultat réside dans une stratégie particulière mise en évidence par Fijalkow (1989) : l'auto-langage. L'auto-langage est défini comme un phénomène qui "consiste en vocalisations à mi-voix, murmures, voire simples mouvements des lèvres". C'est un langage "pour le sujet lui-même à des fins réflexives" qui se "manifeste quand il rencontre des difficultés cognitives" (p.84-85). Il est assimilable à une activité d'auto-répétition dans le sens où son but est de conserver l'information en mémoire mais en diffère car l'information n'est pas transférée en mémoire à long terme, mais plutôt réutilisée lors d'un travail réflexif immédiat. Au cours de cette tâche, nous avons observé que certains enfants bilingues recouraient à cette stratégie : ils répétaient l'item proposé plusieurs fois et donnaient leur réponse à la fin des répétitions. Même si l'auto-langage n'est pas spécifique aux enfants bilingues, il n'a pas été observé chez les enfants monolingues durant la passation. Compte tenu de la difficulté de la tâche, il est possible que l'auto-langage ait permis aux enfants de retenir l'information en mémoire afin d'y réfléchir. L'auto-langage constitue, à notre sens, une manifestation observable des capacités de type métaphonologique car l'enfant fait preuve de flexibilité cognitive en mobilisant pleinement et intentionnellement son savoir langagier pour résoudre la tâche.

La seconde hypothèse concernait la sensibilité des enfants aux segments phonologiques subsyllabiques, examinée au moyen de la tâche de préférence. Les enfants des deux groupes ne se sont pas comportés de façon analogue, ils manifestent une sensibilité légèrement

différente pour certains segments phonologiques. Toutefois une tendance générale ressort de ces résultats : tous les enfants ont une préférence plus marquée pour les grandes unités phonologiques (appelées aussi unités larges) que pour les petites. Les grandes unités, groupes phonémiques composés de deux phonèmes dans la présente étude, sont la rime (-VC) et le début (CV-). Au contraire, les petites unités, composées d'un seul phonème, sont l'attaque (C--), la voyelle (-V-) et le coda (--C). Cette préférence peut être expliquée par la recherche de Liberman, Cooper, Shankweiler & Studdert-Kennedy (1967). En effet, ces auteurs soulignent que les unités larges correspondent à des segments consonantiques discrets du courant acoustique et sont donc des unités d'articulation perçues naturellement par l'auditeur de langue française. En revanche, les petites unités constituent des bruits qui n'évoquent pas la parole. C'est probablement à cause de cette difficulté de perception phonémique que tous les enfants montrent une plus grande sensibilité à la rime et au début. Cependant, les résultats des enfants monolingues divergent de ceux des enfants bilingues quant à l'unité subsyllabique la plus saillante. Effectivement, les enfants monolingues sont plus sensibles à la rime alors que les enfants bilingues sont légèrement plus sensibles au début. D'après Bryant, MacLean, Bradley & Crossland (1990), qui utilisent du matériel anglais, la rime est une unité subsyllabique saillante comparativement aux autres segments phonologiques pour les enfants pré-lecteurs anglais. Martinot & Gombert (soumis) confirment l'importance de la rime avec du matériel français. Par conséquent, la rime est un segment subsyllabique saillant pour les enfants pré-lecteurs français et anglais. Bryant & coll. (1990) rapportent également une sensibilité importante pour l'attaque. On s'attendait d'ailleurs à retrouver cette sensibilité chez les enfants bilingues. Or ce n'est pas le cas. Au contraire, les enfants bilingues montrent une plus grande préférence pour le début des items. Cette préférence peut être expliquée par les mêmes facteurs que la rime : le début, composé de deux phonèmes, est une unité subsyllabique large, qui contient une pointe acoustique et correspond à un signal audible. Sa composition en consonne et voyelle permet à l'enfant de le prononcer facilement. Ensuite, d'un point de vue purement linguistique, la prosodie et l'accentuation de chaque langue doivent être prises en compte. L'accent en français porte généralement sur la dernière syllabe d'un mot ou d'un groupe de mots. Au contraire, la place de l'accent en langue anglaise est plus difficilement prévisible mais peut se trouver sur la première syllabe. Dans le cas de mots monosyllabiques, l'accent est porté sur les phonèmes finaux en langue française et sur les phonèmes initiaux en langue anglaise. De cette façon, étant donné que les enfants monolingues ont une préférence pour la rime et les enfants bilingues pour le début, les résultats peuvent être interprétés comme si du matériel français avait été proposé aux enfants monolingues, ce qui est le cas; mais également comme si du matériel anglais avait été proposé aux enfants bilingues puisqu'ils se comportent comme s'ils étaient face à des items anglais, alors que tous les items sont en langue française. Ces résultats mériteraient donc de plus amples investigations.

6. Perspective

Ce travail, à visée purement exploratoire, nécessite d'être étendu. D'une part il serait intéressant de comparer expérimentalement ce type de résultat avec des enfants monolingues anglais et des enfants bilingues sur du matériel anglais (même si les pseudo-mots sont là pour parer ce type de comparaison). Si les enfants bilingues ont un traitement phonologique spécifique, on devrait retrouver le même type de résultat indépendamment du matériel. En revanche, si tout relève de l'influence linguistique de l'anglais et/ou du français, on devrait trouver des résultats différents en fonction du matériel. Il est donc indispensable de faire la part des choses entre ce qui relève des capacités métaphonologiques strictes, ce qui relève des contraintes linguistiques pures et enfin ce qui relève de l'interaction entre les deux. D'autre part, il apparaît pertinent de s'interroger sur les capacités métalinguistiques des enfants bilingues issus d'un enseignement bilingue immersif. Trop peu de travaux ont trait à

l'enseignement bilingue en France alors que ce dernier existe depuis plus d'une dizaine d'années. Il nous semble donc particulièrement judicieux de nous pencher sur cette population bilingue "disponible" dans notre environnement.

Références

Références bibliographiques

- BIALYSTOK, Ellen (1991). *Language processing in bilingual children*. London : Cambridge University Press.
- BLOOMFIELD, Leonard (1933). *Language*. London: Allen and Unwin.
- BRUCK, Maggie & GENESEE, Fred (1995). Phonological awareness in young second language learners, *Journal of Child Language*, 22: 307-324.
- BRYANT, Peter, MACLEAN, Morag, BRADLEY, Lynette & CROSSLAND, James (1990). Rhyme and Alliteration, Phoneme Detection and Learning to Read, *Developmental Psychology*, 26: 429-438.
- CAMPBELL, Ruth & SAIS, Efsia (1995). Accelerated metalinguistic (phonological) awareness in bilingual children, *British Journal of Developmental Psychology*, 13: 61-68.
- CONTENT, Alain, KOLINSKY, Régine, MORAIS, José & BERTELSON, Paul (1986). Phonetic segmentation in pre-readers: Effects of corrective information, *Journal of Experimental Child Psychology*, 42: 49-72.
- DESHAYS, Elisabeth (1998). *L'enfant bilingue. Parler deux langues : une chance pour votre enfant*. Paris : Robert Laffont.
- FIJALKOW, Jacques (1989). Auto-langage et apprentissage de la lecture, *Enfance*, 42 (1-2): 83-90.
- GOMBERT, Jean-Emile (1990). *Le développement métalinguistique*. Paris: PUF.
- GOMBERT, Jean-Emile & COLE, Pascale (2000). Activités métalinguistiques, lecture et illettrisme, in *L'acquisition du Langage. Le langage en développement au-delà de trois ans*, Kail, Michèle & Fayol, Michel (Eds). Paris: PUF, 117-150.
- LIBERMAN, Alvin, COOPER, Franklin, SHANKWEILER, Donald & STUDDERT-KENNEDY, Michael (1967). Perception of speech code, *Psychological Review*, 74: 431-461.
- MACNAMARA, John (1967). The bilingual's performance: A psychological overview, *Journal of Social Issues*, 23: 58-77.
- MANN, Virginia & WIMMER, Heinz (2002). Phoneme awareness and pathways into literacy: A comparison of German and American children, *Reading and Writing: An Interdisciplinary Journal*, 15: 653-682.
- MARTINOT, Clara & GOMBERT, Jean-Emile (1996). Le développement et le contrôle des connaissances phonologiques chez les enfants pré-lecteurs, *Revue Française de Neuropsychologie*, 6: 251-270.
- MARTINOT, Clara & GOMBERT, Jean-Emile (1998). Les unités phonologiques. Communication orale dans le cadre d'un Symposium organisé par le Pr Michel Fayol pour le XXVI Congrès Européen de Psychologie du Développement, Rennes.
- MARTINOT, Clara & GOMBERT, Jean-Emile (soumis). Implicit onset/rime knowledge in prereaders and beginning readers, *European Journal of Developmental Psychology*.
- PERREGAUX, Christiane (1994). *Les enfants à deux voix: Des effets du bilinguisme sur l'apprentissage de la lecture*. Berlin: Peter Lang.
- ROSNER, Jerome & SIMON, Dorothea (1971). The auditory analysis test: an initial report, *Journal of learning disabilities*, 4: 40-48.
- VYGOTSKY, Lev Semenovitch (1934, 1985). *Pensée et langage*. Paris: Messidor.
- YOPP, Hallie Kay (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly*, 23: 160-177.

LE MARQUAGE ASPECTUEL DE LA SIMULTANÉITÉ CHEZ DES APPRENANTS QUASI-NATIFS FRANCOPHONES DE L'ANGLAIS DANS UNE TÂCHE NARRATIVE

Pascale LECLERCQ

ppeleclercq@free.fr

UMR 7023, Université Paris 8 Vincennes - Saint Denis

2 rue de la Liberté

93526 Saint-Denis Cedex

Abstract : This comparative study on the expression of “ongoingness” across languages investigates how aspectual marking is used in narratives by French and English native speakers, and near-native French learners of English, to express simultaneity.

Notre étude part du constat que les apprenants d'une L2, même très avancés, peinent à atteindre le niveau de compétence des locuteurs natifs. Nous postulons que cela est lié au fait que les différentes possibilités grammaticales offertes par les langues influencent le locuteur d'une langue donnée au moment de la sélection de l'information en vue de la verbalisation (Levelt, 1989, Slobin, 2003). Nous nous intéressons donc à l'interface entre la perception du monde et la verbalisation. En fonction des marqueurs grammaticaux présents dans leur langue, les locuteurs vont focaliser leur attention sur tel ou tel élément de la réalité. La verbalisation du monde diffère donc en fonction de la langue du locuteur, et des formes linguistiques offertes par celle-ci (Slobin, 1996, von Stutterheim, 2003).

Pour tester cette hypothèse et voir dans quelle mesure les apprenants intègrent les moyens offerts par la langue cible, nous avons mis au point un dispositif expérimental qui suscite l'expression de l'aspect « en déroulement ».

Le choix de l'aspect est pertinent pour la comparaison du français et de l'anglais, car l'anglais code l'aspect « en déroulement » morphologiquement sur le verbe, alors que le français a recours soit au présent simple, soit à la périphrase « en train de » pour l'exprimer. Nous supposons que ces divergences grammaticales ont des conséquences sur les choix opérés au niveau conceptuel.

Notre analyse s'insère dans un projet européen sur l'acquisition des langues étrangères par des apprenants de niveau avancé (Carroll, Lambert, Natale, Starren, von Stutterheim, à paraître). Les mêmes supports expérimentaux sont utilisés à Heidelberg (Allemagne) et à Nijmegen (Pays Bas) dans le cadre de ce projet pour comparer d'autres paires de langues qui présentent un contraste au niveau du marquage aspectuel.

Nous utilisons deux tâches expérimentales dans notre travail de thèse. Les résultats de la première expérience – une tâche de compte-rendu *online* de vidéo clips suscitant l'expression de l'aspect « en déroulement » – nous ont montré que « en train de » est employé par les francophones à hauteur de 26% (contre 91% pour *be + V-ing* chez les anglophones), et que cette périphrase n'est utilisée que pour rendre compte de certains types de situations : on ne la trouve jamais pour décrire des situations de mouvement ou de déplacement, alors qu'elle est fréquente avec les scènes montrant une activité orientée vers un but, ainsi qu'avec les situations de simultanéité. Nous souhaitons donc affiner notre analyse de l'utilisation de l'aspect « en déroulement » par les francophones apprenants de l'anglais et nous avons mis sur pied un nouveau dispositif expérimental, constitué de clips publicitaires suscitant l'expression de la simultanéité, ce qui permet de tester les marqueurs adoptés par les

locuteurs francophones et anglophones natifs, et par les apprenants francophones de l'anglais, lorsqu'il s'agit de faire un récit d'événements simultanés. Nous posons donc les questions suivantes :

- Quels moyens linguistiques sont employés par les trois groupes de locuteurs (FrL1, AngL1, FrL1 AngL2) pour exprimer la simultanéité ?
- Les préférences sont-elles divergentes selon les langues sources ?
- Le marquage aspectuel joue-t-il un rôle structurant pour l'expression de la simultanéité ?

Nous cherchons désormais comment les locuteurs francophones et anglophones et les apprenants francophones de l'anglais expriment la simultanéité : à l'aide de temporelles (« en même temps que », *while...*), et/ou à l'aide du marquage aspectuel (« en train de », *Ving*). Ex.: “*He is licking off the ketchup, and at the same time, his brother walks into the room.*” (Locuteur anglophone natif).

Ainsi on pourra voir dans quelle mesure les formes aspectuelles sont utilisées par les locuteurs anglophones et francophones, et par les apprenants avancés francophones de l'anglais, pour l'expression de la simultanéité.

1. Méthodologie

Nous avons choisi d'avoir recours à une tâche de récit pour l'étude de la simultanéité. Dans la tâche de compte-rendu *online* de notre première expérience, les locuteurs n'avaient en effet pas le temps de planifier la verbalisation. Avec une tâche de récit au contraire les locuteurs peuvent planifier la macrostructure de leur narration : nous souhaitons donc voir si le marquage aspectuel joue un rôle de structuration de l'information dans le récit.

1.1. Présentation de la tâche

Notre tâche de récit est composée de 5 clips vidéo publicitaires, sans paroles ou avec un message verbal minimal et non nécessaire à la compréhension de la situation. Ces cinq clips sont intitulés Canal +, Saumon, Hotdog, Soupe et Supermarché. Les sujets visionnent les publicités autant de fois qu'ils le souhaitent, afin de mémoriser la séquence d'événements. Après avoir visionné une publicité, le sujet en effectue un récit, puis visionne la publicité suivante et ainsi de suite.

1.2. Corpus

Pour cette tâche nous disposons de 12 locuteurs francophones natifs, de 20 locuteurs anglophones natifs, et de 10 apprenants quasi-natifs francophones de l'anglais.

Tous ces locuteurs sont diplômés de l'université, ce qui garantit la comparabilité des résultats ; les apprenants quasi-natifs francophones de l'anglais sont des enseignants certifiés ou agrégés d'anglais. Les enregistrements ont été réalisés chez les sujets dans des conditions informelles.

1.3. La simultanéité : une définition

Les situations présentées dans les clips du support ont été sélectionnées car elles présentent des événements qui se déroulent de manière simultanée. Toutes ces situations présentent deux protagonistes ou plus.

Il y a simultanéité lorsque deux (ou plus) événements prennent place dans le même intervalle temporel. La simultanéité est perceptible lorsque ces deux événements se déroulent dans un espace commun (au moins aux yeux de l'énonciateur). Les événements simultanés peuvent n'avoir aucun lien entre eux ou être entremêlés. Nous nous intéressons donc à l'interface entre la perception de deux événements simultanés par un locuteur et à la manière dont il va

représenter verbalement cette situation, en fonction des moyens grammaticaux/linguistiques qui sont à sa disposition.

Les cinq clips de notre étude présentent deux événements (E1, événement cadre, et E2) ayant lieu de manière simultanée. Certains événements sont itérés : c'est le cas dans les clips Soupe et Saumon. Dans certains clips, E1 et E2 sont deux événements qui se déroulent en parallèle, sans que l'un ne vienne perturber ou interrompre le déroulement de l'autre. C'est le cas pour le clip « Saumon » où l'action du cuisinier qui prépare le repas (E1) et celle du chat qui dérobo des tranches de saumons (E2) se déroulent en parallèle, sans que E1 n'interrompe E2. Dans ce cas, on n'a pas de rupture événementielle.

Dans d'autres cas, un événement E2 vient interrompre un autre événement E1 : soit E2 met un terme à E1, comme c'est le cas dans le clip Hotdog où le jeune garçon qui lèche le poster (E1) est interrompu par l'irruption dans la chambre de son grand frère (E2), soit les deux événements se déroulent ensuite en parallèle, comme dans le clip Canal+ où l'homme continue de regarder la télévision (E1) après l'irruption des mécaniciens qui s'activent autour de son canapé (E2). On parle, pour ces clips, de rupture événementielle. Dans le clip Canal+, E2 est inclus dans E1 : les événements sont « emboîtés ». On peut donc résumer les différents traits caractéristiques des clips de la manière suivante :

1. Pas de rupture événementielle : E1 // E2
2. Rupture événementielle (E2 interrompt E1) ; on peut alors distinguer trois cas :
 1. soit E2 met un terme à E1
 2. soit les deux événements continuent à se dérouler de manière parallèle, sans que le terme de l'événement ne soit atteint
 3. Inclusion (E2 est inclus dans E1 : les événements sont « emboîtés »)
3. Itération (E2 est un événement complexe composé d'événements répétés)

Nous souhaitons évaluer l'impact de ces différents traits sur les choix des locuteurs au moment de la verbalisation.

Clips	Rupture événementielle	E2 met un terme à E1	E2 et E1 se superposent	E2 inclus dans E1	Itération
Canal +	+	-	+	+	-
Saumon	-	-	+	-	+
Hotdog	+	+	-	-	-
Soupe	-	-	+	+	+
Supermarché	-	-	+	-	-

Tableau 1 - Caractéristiques des clips

Dans cette analyse, nous nous intéresserons particulièrement au trait « rupture événementielle » : nous tenterons de déterminer si ce trait déclenche l'utilisation du contraste aspectuel dans les récits des trois groupes de locuteurs.

1.4. Exemples de récits

- Récit oral FrL1, Clip Hotdog, sujet MAA

Alors on voit un garçon qui rentre de l'école, qui se précipite dans la cuisine et qui va ouvrir son frigidaire, prendre un pot de moutarde et un pot de ketchup, ensuite il se fait un sandwich très rapidement et il monte dans sa chambre en courant et en fait il dévore son sandwich et en fait ça explose sur un poster d'une fille à moitié nue dans sa chambre, et donc en fait après il se met en fait à lécher la moutarde qui a explosé sur le ventre de la fille, et juste à ce moment là en fait il y a un garçon qui doit être sûrement son grand frère qui rentre dans sa chambre, et qui lui fait un clin d'œil.

- Récit oral AngL1, Clip Hotdog, sujetS22

So a young boy comes home, maybe about fourteen, and he throws his bag on the floor. And then he goes into the kitchen and he finds some hotdogs, some ketchup, some mustard, and he makes a hotdog. He goes upstairs to his bedroom, and while he is eating his hotdog, he squirts some of the mustard and ketchup onto a poster on his bedroom wall. He goes to lick off the ketchup, and the mustard, and while he is doing this, his brother walks into the room and thinks he is licking the poster and starts to tease him.

• Récit oral FrL1 AngL2, Clip Hotdog, sujet ETI

Ok, so there's a young boy coming back from school, a young teenager, throws his bag as he goes into the house, goes into the kitchen, takes some ketchup or either mustard or mayonnaise, I can't remember, and he makes a hotdog, loads of mayonnaise, loads of ketchup, then he goes into his room, starts eating it, bites it, and some of the sauce spurts on some kind of poster of a sexy woman half naked, and then he goes to the poster and starts licking it, to lick the sauce, and as he's licking the sauce his brother, well, his brother I think comes in and makes fun of him, and he's all shy, and he didn't expect that basically.

1.5. Caractéristiques des récits

1.5.1. Les adverbiaux temporels

Tous les récits contiennent des adverbiaux temporels, qui permettent de marquer la concomitance à travers la rupture temporelle (« tout d'un coup »), l'ouverture d'un cadre temporel au sein duquel prennent place une série d'événements (« pendant que »), ou la successivité (« et puis »). Dans ce dernier cas, le contexte (notamment le marquage aspectuel) nous permet de déduire s'il s'agit de successivité ou de simultanéité.

Adverbiaux	Fr L1 N=12	Ang L1 N=20	Fr L1 Ang L2 N=10
Rupture temporelle (%)	33,2	17,1	14
Intervalles parallèles (%)	15	18,3	28
Séquentialité (%)	23,4	39,4	42
Juxtaposition (%)	11,7	7,2	8
Pas de simultanéité (%)	1,7	/	/
Itérativité (%)	8,2	8,3	4
Autres (%)	6,7	6,7	4

Tableau 2 - Les compléments adverbiaux temporels (5 clips)

Ces chiffres révèlent une nette préférence des anglophones pour l'expression de la séquentialité, alors que les francophones privilégient l'expression de la rupture temporelle. Quant aux apprenants, comme les anglophones, ils emploient majoritairement les adverbes exprimant la séquentialité, toutefois ils emploient les compléments adverbiaux marquant l'ouverture de deux intervalles parallèles (« at the same time », « while ») de manière beaucoup plus importante que les deux groupes de locuteurs natifs.

Nous verrons comment les adverbiaux temporels sont employés en complément du marquage aspectuel.

1.5.2. Temps et aspects

Dans l'immense majorité des récits recueillis, les verbes sont au présent. Toutefois on observe l'emploi de formes aspectuelles à certains endroits du récit. On distingue le marquage aspectuel morphologique, comme *V-ing* en anglais, du marquage lexical de l'aspect (« en train de ») en français. Dans cette langue, c'est le présent simple qui est employé : l'utilisation de « en train de » dans un récit peut alors avoir une valeur contrastive. En anglais, le temps du récit est, selon les clips, le présent simple (clip Hotdog) ou le présent *be + V-ing*. Ainsi, dans les récits du clip Hotdog, au présent, l'emploi de *V-ing* sera interprété comme contrastif. Au contraire, dans les récits du clip Supermarché,

majoritairement à la forme *be + V-ing*, c'est le présent simple qui recevra une lecture contrastive.

L'utilisation du contraste aspectuel à des fins de structuration du récit a déjà été montrée dans de nombreuses études, comme celles de Hopper (1979) sur le russe, de Givón et Bickerton sur les créoles hawaïien et guyanais (cités par Wenzell, 1989), de Lambert, Carroll & von Stutterheim (2003), de Chuquet (1994), Schmiedtova (2004), etc. On peut donc se demander si le contraste aspectuel entre formes non marquées (présent simple) et formes aspectuelles (*be + V-ing*, « en train de ») fonctionne de la même manière en français et en anglais pour rendre concomitants les événements E1 et E2.

L'analyse de notre corpus nous donne les résultats suivants pour le contraste aspectuel :

	Hotdog	Supermarché	Canal +	Saumon	Soupe
Fr L1	25%	42%	42%	50%	58%
Ang L1	65%	47%	72%	37%	55%
Fr L1Ang L2	90%	60%	70%	40%	40%

Tableau 3 - Le contraste aspectuel : résultats

2. Analyse des résultats

2.1. Résultats Fr L1

En français, on relève des formes progressives contrastives à hauteur de 50% dans quatre des cinq clips de notre étude. Les récits des francophones sont toujours au présent simple, l'emploi de « en train de » est donc nécessairement contrastif. Ce chiffre nous semble très élevé : en effet, nous avons trouvé dans une première expérience de compte-rendu *online* de vidéo clips suscitant l'emploi des formes progressives, un taux de « en train de » de 26%, contre 91% de formes en *V-ing* chez les anglophones. Il semble donc que la fréquence d'utilisation de cette périphrase soit faible. Le taux de 50% relevé dans notre corpus pourrait donc confirmer notre hypothèse de départ selon laquelle la simultanéité suscite l'emploi de « en train de ».

Parmi les quatre clips pour lesquels « en train de » est fortement utilisé, on relève trois clips sans rupture événementielle (Saumon, Soupe et Supermarché), et un clip avec rupture événementielle (Canal+). Quant au clip Hotdog, qui suscite peu l'emploi de « en train de », il comporte le trait rupture événementielle.

Le critère « rupture événementielle » ne semble donc pas être déterminant en français pour susciter l'emploi de « en train de ». Par contre, le fait d'avoir E1 et E2 superposés, se déroulant en parallèle, comme c'est le cas pour Saumon, Soupe, Supermarché et Canal+, semble entraîner l'emploi du contraste aspectuel chez les francophones (voir tableau 1).

On peut néanmoins se demander pourquoi le taux de formes progressives est plus faible dans le clip Hotdog. Comme le fait remarquer Chuquet (1994), une succession d'événements brefs envisagés dans leur globalité est généralement rendue par le présent simple ; les événements présentés dans ce clip étant très courts (le jeune garçon a à peine commencé à lécher le poster lorsque son grand frère entre dans la chambre), il est donc logique d'employer le présent et non le progressif qui envisage les événements sous l'angle de leur déroulement. On a vu dans l'analyse du corpus Clips que le critère « durée » de la scène influe sur l'utilisation de « en train de ». Ainsi le clip Hotdog, qui présente une succession rapide d'événements brefs, n'invite-t-il pas à l'utilisation de « en train de ».

Cela ne semble pas être le cas en anglais : nous verrons dans le paragraphe suivant que les anglophones natifs utilisent *V-ing* de manière contrastive à hauteur de 65% dans cette scène.

Nos résultats confirment ceux trouvés à partir de notre première tâche expérimentale de compte-rendu *online* : « en train de » semble largement employé dans les situations de simultanéité, tout particulièrement lorsqu'il s'agit d'activités orientées vers un but (préparer

le repas, remplir le caddy) et d'événements ayant une certaine durée. Il s'agit de voir dans quelle mesure la faible utilisation de « en train de » pour signaler la rupture événementielle, comme c'est le cas dans le clip Hotdog, est compensée par un recours important aux adverbes de rupture temporelle.

2.2. Résultats Ang L1

Chez les anglophones, tous les récits sont à la forme en *V-ing* sauf dans Hotdog : en effet, ce clip présente une série d'actions rapides se succédant les unes aux autres : le présent simple en combinaison avec des verbes à deux états (de type « *enter* », dont le sémantisme implique à la fois le début et la fin de l'événement) est donc préféré. Pour ce clip, *V-ing* est utilisé de manière contrastive à hauteur de 60%. Dans les quatre autres clips, on trouve majoritairement le présent progressif, et c'est le présent simple qui est employé pour marquer le contraste.

On observe que les clips qui ne présentent pas de rupture événementielle (Supermarché, Saumon, Soupe) ont un contraste aspectuel plus faible (respectivement 47, 37 et 55%). Par contre, les clips avec rupture événementielle (Canal +, Hotdog) présentent un contraste aspectuel très fort (respectivement 72 et 65%).

Le critère de rupture événementielle semble donc pertinent pour les anglophones mais pas pour les francophones. Il s'agit de voir comment les adverbes sont employés par ces sujets en combinaison avec le contraste aspectuel pour les deux clips avec rupture événementielle.

2.3. Résultats Fr L1 Ang L2

Les apprenants suivent de très près le schéma de marquage aspectuel des anglophones : dans les clips Saumon et Canal +, on relève des taux de contraste aspectuel identiques chez les anglophones et les apprenants. Dans les clips Supermarché et Hotdog, ils dépassent même le taux de contraste aspectuel relevé chez les anglophones natifs, même si dans le clip Soupe ils emploient le contraste aspectuel de manière légèrement inférieure aux taux observés chez les deux groupes de locuteurs natifs. Dans tous les cas, les apprenants s'écartent largement de la norme des locuteurs francophones natifs, et utilisent le contraste aspectuel de manière très similaire aux anglophones. Notre groupe d'apprenants quasi-natifs semble non influencé par les usages de Fr L1 pour le marquage du progressif.

2.4. Compléments adverbiaux dans les Clips présentant une rupture événementielle (Hotdog, Canal+)

Nous allons maintenant observer ce qui se passe pour les trois groupes de locuteurs en ce qui concerne l'emploi des compléments adverbiaux dans les deux clips qui présentent une rupture événementielle : Hotdog et Canal+. On a vu que les francophones emploient peu « en train de » avec Hotdog, mais qu'ils l'utilisent beaucoup avec le clip Canal+. Il s'agit désormais d'analyser l'usage qui est fait des adverbes de rupture temporelle dans les récits de ces deux supports.

Adverbiaux	Fr L1 N=12	Ang L1 N=20	Fr L1 Ang L2 N=10
Rupture temporelle	66,5%	20%	60%
Intervalles parallèles	8,5%	25%	30%
Séquentialité		40%	10%
Juxtaposition	25%	10%	
Pas de simultanéité		5%	

Tableau 4 - Compléments adverbiaux, clip Hotdog

En ce qui concerne le clip Hotdog, on observe que tous les énoncés contiennent des adverbiaux temporels ; les francophones privilégient l'emploi des adverbes de rupture

temporelle (« et là », « à ce moment-là », « tout à coup »...) alors que les anglophones ont plutôt recours à la séquentialité (« then »). En revanche, les apprenants francophones de l'anglais choisissent également les adverbiaux de rupture temporelle, dans les mêmes proportions que les autres francophones (60%/66,5%).

Adverbiaux	Fr L1 N=12	Ang L1 N=20	Fr L1 Ang L2 N=10
Rupture temporelle	58%	40%	10%
Intervalles parallèles	16,5%	20%	30%
Séquentialité	8,5%	30%	40%
Juxtaposition	8,5%		20%
Pas de simultanéité	8,5%	10%	

Tableau 5 - Compléments adverbiaux, clip Canal+

Dans le clip Canal+, les francophones préfèrent largement (58%) les adverbes exprimant la rupture temporelle, tout comme les anglophones (40%), qui marquent néanmoins beaucoup plus la séquentialité (30%). Quant aux apprenants, ils se détachent des schémas suivis par les deux groupes de locuteurs natifs en n'ayant presque jamais recours aux adverbes de rupture temporelle, et privilégient l'expression de la séquentialité et des intervalles parallèles.

On observe donc que les francophones, qu'ils aient ou non recours au contraste aspectuel pour marquer la rupture événementielle (clips Canal+ et Hotdog), emploient très fortement les adverbes de rupture temporelle (Hotdog : 66,5%, Canal+ : 58%).

Les anglophones qui marquent très fortement le contraste aspectuel dans ces deux clips, n'ont pas systématiquement recours aux adverbes de rupture temporelle (la séquentialité l'emporte avec 40% dans le clip Hotdog). On observe en effet que les anglophones privilégient les adverbes exprimant la séquentialité, en combinaison avec un contraste aspectuel fort.

Quant aux apprenants, ils suivent les schémas des locuteurs anglophones pour le marquage aspectuel, mais utilisent les adverbes temporels soit comme les francophones (Hotdog) soit d'une manière qui leur est propre (Canal+), en n'employant pas d'adverbes de rupture temporelle.

3. Conclusion

Les anglophones marquent la rupture événementielle avant tout par le contraste aspectuel, en combinaison avec des adverbes exprimant la séquentialité. Les francophones au contraire emploient les adverbes de rupture temporelle, avec ou sans contraste aspectuel. Il semble donc que la forme en *V-ing*, grammaticalisée, soit bel et bien employée pour structurer les récits des anglophones lorsqu'il s'agit de rendre compte d'événements simultanés ; « en train de » semble aussi jouer un rôle structurant dans ces récits pour les situations se déroulant en parallèle, qu'il y ait ou non rupture événementielle.

Les apprenants très avancés francophones de l'anglais se détachent nettement du schéma francophone de marquage aspectuel, et tendent à se rapprocher du modèle anglophone. Par contre le bilan est plus nuancé en ce qui concerne l'emploi des adverbes temporels : les apprenants continuent parfois à les employer de manière francophone (clip Hotdog), alors que dans d'autres circonstances (clip Canal+), les choix des apprenants divergent de ceux des deux groupes de locuteurs natifs, ce qui semble plaider en faveur d'une interlangue des apprenants avancés. Si le marquage aspectuel ne leur pose pas de réel problème, les apprenants très avancés semblent néanmoins encore influencés par leur langue maternelle en ce qui concerne la structuration des récits. Ce domaine semble donc très résistant aux remaniements conceptuels opérés par les locuteurs d'une L2, même très avancés.

Références bibliographiques

- CARROLL, Mary, LAMBERT, Monique, NATALE, Silvie, STARREN, Marianne & VON STUTTERHEIM, Christiane (à paraître). Being specific: the role of aspect in event construal when grounding events in context. A cross-linguistic comparison of advanced second language learners (L1 French–L2 English, L1 German–L2 English, L1 Dutch–L2 French, L1 Italian–L2 German), in *Processes and Outcomes. Explaining Achievement in Language Learning*, Habertzettl, Stephanie (Ed). Berlin: de Gruyter.
- CHUQUET, Hélène (1994). *Le présent de narration en anglais et en français. Linguistique contrastive et traduction*. Paris : Ophrys.
- LAMBERT, Monique, CARROLL, Mary, VON STUTTERHEIM, Christiane (2003). La subordination dans les récits d'apprenants avancés francophones et germanophones de l'anglais, *AILE 19 Les énoncés complexes et leur développement dans l'acquisition des langues, Aile-Encrages*, 41-69.
- LEVELT, Willem J.M. (1989). *Speaking: from intention to articulation*. Cambridge, Massachussets: MIT Press.
- SCHMIEDTOVA, Barbara (2004). *At the same time... The expression of simultaneity in learner varieties*. Thèse de doctorat. MPI Series in Psycholinguistics.
- SLOBIN, Dan I., (1996). From 'Thought to Language' to 'Thinking for Speaking' in *Rethinking Linguistic Relativity*, Gumperz, John & Levinson, Stephen (Eds). Cambridge, Cambridge University Press, 70-96.
- SLOBIN, Dan I. (2003). Language and thought online: Cognitive consequences of linguistic relativity, in *Language in mind: Advances in the investigation of language and thought*, Gentner, Dedre & Goldin-Meadow, Susan (Eds). Cambridge, Massachussets: MIT Press, 157-191.
- VON STUTTERHEIM, Christiane (2003). Linguistic Structure and information organisation: The case of very advanced learners, in *EUROSLA Yearbook*, vol. 3, Foster-Cohen, Susan & Pekarek Doehler, Simona (Eds). Amsterdam: John Benjamins Publishing Company, 183-206.
- WENZELL, Vanessa E. (1989). Transfer of Aspect in the English Oral Narratives of Native Russian Speakers, in *Transfer in Language Production*, Chapter 5, Dechert, Hans & Raupach, Manfred (Eds). Norwood, N.J. (New Jersey): Ablex Publishing Corporation, 71-97.

LE RÔLE DE LA CONVERSATION DANS LA TRANSITION DES ÉNONCÉS À UN MOT VERS LES ÉNONCÉS À DEUX MOTS : UNE ÉTUDE EXPÉRIMENTALE

Silvia NIEVA

silnieva@gmail.com

EQUIAL-Equipo de Adquisición del Lenguaje

Facultad de Psicología (Campus de Somosaguas)

28223 Madrid

Espagne

Abstract : The transition in the production from single-word utterances to multiword utterances depends upon the conversational sequences that characterize adult-child exchanges. This experimental study aims at facilitating this transition. The results show the diversity of such exchanges and the key role which they play in first language development acquisition.

1. Introduction

Ce projet s'intègre dans un modèle général de type constructiviste-émergentiste. Plusieurs études s'inscrivant dans cette approche ont montré que l'acquisition des structures grammaticales part de la capacité humaine à apprendre (Bates & MacWhinney, 1987; Elman, 1993, 1999; Elman, Bates, Johnson, Karmiloff-Smith, Parisi & Plunkett, 1996; Karmiloff-Smith, 1992; MacWhinney, 1999). Dans cette approche, l'enfant extrait l'information sur les structures linguistiques de sa langue à partir de l'interaction avec l'adulte (input) d'une part, et de l'élaboration interne de cette information d'autre part.

Ces interactions enfant-adulte ont très souvent une structure conversationnelle. Le modèle conversationnel fut proposé initialement pour la conversation adulte (Sacks, Schegloff & Jefferson, 1974). D'autres auteurs dans une approche plus centrée sur le développement tiennent compte du langage adapté des parents résultant des capacités précoces de leur bébé à interagir (Dore, 1985; Snow, 1977; Veneziano, 1989). Situé dans cette orientation, notre projet analyse non seulement les productions que la mère adresse à l'enfant afin de lui demander plus d'information, mais aussi les réactions de l'enfant à celles-ci.

Ainsi, les très jeunes enfants dont les productions ne portent que sur des mots isolés et des amalgames (qui correspondent à une seule unité de stockage) peuvent se servir des éléments fournis par l'adulte au cours de la conversation. En portant sur un même événement-situation, ces enfants ajouteraient un deuxième élément à leur tour de parole qui est relié sémantiquement à leur première production spontanée (Veneziano, 1999).

Divers travaux témoignent de ce type d'émissions, et différentes terminologies ont été proposées, dont "holistic successive single-word utterances" (SSWUs) par Bloom (1973), "sequences" par Greenfield & Smith (1976), "propositions across utterances" par Ochs, Schieffelin & Platt (1979) et constructions verticales par Scollon (1979). En partant des données longitudinales, Veneziano (1999) propose une classification des SSWUs, en français : des SEMeS (Suite d'Énoncés à un Mot en Succession), qui repose principalement sur le niveau de dépendance vis-à-vis des enchaînements conversationnels. Ce travail est une adaptation empirique de cette classification et une reproduction de ces mécanismes interactifs d'une manière contrôlée dans une situation expérimentale.

On ne trouve pas beaucoup d'études expérimentales qui visent à tester le rôle des SEMeS sur l'émergence de la combinaison de mots. Schwartz, Chapman, Prelock, Terrell, et Rowan (1985) ont construit des situations qui élicitent chez l'enfant la production de deux mots et ils ont observé des changements significatifs dans le post-test. Dans notre étude expérimentale, on vise à observer si, en fournissant de manière intensive à l'enfant des opportunités pour produire ces SEMeS, il progresse plus rapidement dans ce développement.

Nous essayerons d'amener l'enfant à mettre en relation la succession des tours de parole dans une conversation et ainsi réunir deux productions à un mot ou amalgame au sein d'une séquence.

2. Transcription et codage des données.

Toutes les données recueillies pour cette étude ont été transcrites orthographiquement d'après les standards CHAT du réseau international Child Language Data Exchange System (CHILDES, Mac Whinney, 2000). Ces transcriptions ont été réalisées sur 30 minutes consécutives d'enregistrement de langage dans des conditions naturelles et elles incluent des indications de situation et de contexte. L'utilisation de ce système standard permet le partage des données codées avec plusieurs chercheurs.

Comme unité d'analyse, nous avons choisi les productions de deux ou plusieurs mots successifs se référant à un événement (Veneziano, Sinclair et Berthoud, 1990). On considère l'événement en tant que séquence conversationnelle autour d'une fonction communicative.

Pour la codification, nous avons utilisé un système de classification en niveaux (Veneziano, 1999). On tient compte des différentes structures transitionnelles entre la production des énoncés à un mot simple et ceux qui sont à plusieurs mots (cf. tableau 1) :

- Le type de production :

1W : production répétée d'un seul mot¹ pendant deux tours interrompus par l'intervention de l'adulte qui sollicite un deuxième mot différent.

Amalgames : ce sont les formules ou structures "roted-learned" (Lieven, Pine et Baldwin, 1997). Les amalgames sont des productions dont les composants n'apparaissent pas isolément dans le vocabulaire de l'enfant et pour lesquelles il n'existe pas de régularités positionnelles dans des productions différentes (Vihman et McCune, 1994). Ces variantes sont traitées dans la codification comme productions à un mot simple.

SEMeS : Suite des Émissions à deux Mots en Succession. Production de deux ou plusieurs mots différents interrompus par l'intervention de l'adulte, ou par une pause de plus de 0.5 secondes.

2W : production d'énoncés à deux mots ou plus.

- Le type de SEMeS :

Inter-tours : les mots produits par l'enfant sont séparés par l'intervention de l'adulte ou par une pause de plus de 2 secondes.

Intra-tour : les mots produits par l'enfant sont séparés par une pause d'entre 0.5 et 2 secondes.

- SEMeS – Le niveau de la conversation :

Cg : SEMeS générées par la conversation. L'enfant produit le deuxième mot motivé par l'intervention antérieure de l'adulte.

Ci : SEMeS indépendantes de la conversation. Le deuxième mot produit par l'enfant n'est pas élicité par l'intervention de l'adulte.

¹ On se réfère à mot comme unité, soit mot ou amalgame.

- SEMeS Cg-SEMeS intra – L’intervention du partenaire :

Sol : l’adulte sollicite par une question la production d’un deuxième mot.

Le type de question que pose l’adulte dépend de la catégorie fonctionnelle du deuxième mot que l’on sollicite en fonction de variables contextuelles.

Imit : la production du deuxième mot de l’enfant est le résultat de l’imitation partielle ou complète du contenu du tour antérieur de l’adulte.

Type de production				
1W				
Amalgame				
SEMeS	Type de SEMeS	Niveau de la conversation (intervention du partenaire)		Niveau du discours (unité thématique)
	Inter-tours	cg (SEMeS générées par la conversation)	imit Modèle pour l'imitation	-
			sol Réponse sollicitée	-
		ci (SEMeS indépendantes de la conversation)	-	dd (SEMeS dépendantes de la conversation)
			-	di (SEMeS indépendantes de la conversation)
	Intra-tour	-		dd (SEMeS dépendantes de la conversation)
				di (SEMeS indépendantes de la conversation)
		imit Modèle pour l'imitation		-
2W	Niveau de la conversation (intervention du partenaire)		Niveau du discours (unité thématique)	
	-		dd (SEMeS dépendantes de la conversation)	
			di (SEMeS indépendantes de la conversation)	
	imit Modèle pour l'imitation		-	

Tableau 1 - Classification par niveaux

- Le niveau du discours (unité thématique) :

Cette variable tient compte des mots de la série que produit l’enfant qui ont été introduits préalablement dans le discours.

Dd (SEMeS dépendantes du discours) : Au moins un des deux mots de la série a été produit auparavant par l’enfant ou par l’adulte, dans les dix tours qui précèdent la série.

Di (SEMeS indépendantes du discours) : Aucun des deux mots de la série n'a été produit auparavant par l'enfant ou par l'adulte, dans les dix tours qui précèdent la série.

2.1. Exemples

Exemple n.1. Marylou, 21 mois. Échange réciproque d'un seul mot (1W)
La mère et l'enfant jouent à mettre des perles dans un collier.

*CHI : encore.

*MOT : encore quoi ?

*CHI : encore.

*MOT : encore mettre des perles ?

Exemple n.2. Constantin, 22 mois. SEMeS inter-tours générées par la conversation (modèle pour l'imitation). La mère et l'enfant regardent par la fenêtre.

*CHI : ebafe [: arbre].

*MOT : ah oui # c'est les arbres là # oui.

*MOT : je sais # il y a des arbres sur la terrasse là- haut.

*CHI : cabo [: là-haut]. [+ im]

Exemple n.3. Angèle, 21 mois. SEMeS inter-tours générées par la conversation (sollicitation du deuxième mot par une question). La mère et l'enfant regardent un livre.

*CHI : papa.

*MOT : ça c'est le papa du petit pingouin.

*MOT : qu'est-ce qu'il lui fait son papa ?

*CHI : alin [: câlin].

Exemple n.4. Constantin, 22 mois. SEMeS inter-tours indépendantes de la conversation
L'enfant est à côté de son lit où il y a des jouets.

*MOT : tu veux faire dodo avec ça ?

*CHI : egoudou [: grenouille].

*MOT : oui # c'est la grenouille.

*CHI : edodo [: dormir].

*MOT : dodo ? [: dormir]

*MOT : elle fait dodo la grenouille # tu crois ?

3. Hypothèses

Nous avons construit des situations aptes à poser des questions servant à éliciter, chez l'enfant, un deuxième mot lié sémantiquement à sa production spontanée précédente, ce qui est susceptible d'amener l'enfant à produire des SEMeS générées par la conversation. On s'attend à ce qu'en post-test, les enfants du groupe expérimental montrent un accroissement plus élevé de la production de SEMeS et des énoncés à deux mots que les enfants du groupe contrôle, notamment sur les énoncés que l'enfant produit indépendamment de la conversation et du discours. C'est-à-dire, les deux mots produits par l'enfant, soit dans une série du type SEMeS ou un énoncé à deux mots, n'ont pas été élicités par l'adulte (niveau de la conversation), ni utilisés par les interlocuteurs dans les tours précédant la série (niveau du discours) non plus.

4. Méthode

Dix enfants monolingues francophones ont été sélectionnés. Ils sont âgés entre 19 mois et 25 mois et séparés en deux groupes : expérimental et contrôle. D'après des rapports parentaux, chaque enfant a atteint les principales étapes du développement aux âges normatifs.

L'échantillonnage a été réalisé à partir des réponses données par les parents à l'adaptation française du MacArthur CDI. (Kern, 1999), en tenant compte de la quantité de vocabulaire. On établit plus de 70 mot-types comme niveau de base, en ayant comme premier critère de sélection que, de l'avis des parents, les enfants ne produisent spontanément que des énoncés à un mot.

5. Procédure

L'étude a lieu dans l'environnement naturel de l'enfant. Les capacités de production de chaque enfant sont mesurées au cours de neuf visites (dont une de familiarisation), six séances expérimentales et deux séances d'observation de l'interaction avec la mère (une séance en pré-test et une autre en post-test). Toutes les séances sont enregistrées de manière audiovisuelle.

L'expérience s'effectue sur une période d'environ trois semaines. Les deux séances d'observation de l'interaction naturelle ont une durée de 30 minutes. Les séances ont lieu trois fois par semaine, à l'exception du post-test, qui a lieu une semaine après les situations-test.

La procédure expérimentale est individuelle et a pour but l'élicitation par la conversation d'un deuxième élément à partir d'une première émission spontanée ou élicitée à la suite de la présentation de stimuli visuels. Les séances expérimentales consistent à présenter à chaque enfant un cahier-album contenant des images familières (des photos). Ces images représentent de scènes quotidiennes contenant des référents qui font partie des items lexicaux répertoriés dans l'inventaire de vocabulaire que les parents ont rempli au préalable. Les stimuli sont composés individuellement contenant des éléments de son vocabulaire qui puissent être liés par des relations sémantiques parmi les suivantes (Braine, 1976, Brown, 1973) : objet-possesseur, objet-bénéficiaire, objet/personne-locatif, agent-action, agent-objet.

L'expérimentateur montre les images à l'enfant une par une en lui laissant le temps pour répondre. S'il ne répond pas, l'expérimentateur pointe son doigt vers un objet ou personne (personne ou objet animé) dans l'image et lui pose une question du type: c'est qui/ c'est quoi?, en sollicitant le premier mot de la série.

Ensuite, les types d'interventions de l'expérimentateur visent à solliciter deux variantes de différents niveaux de SEMeS inter-tours générées par la conversation : (1) Les SEMeS qui fournissent à l'enfant un modèle pour l'imitation :

Exemple n.5. SEMeS inter-tours générées par la conversation

Relation sémantique à travailler : Objet + possesseur.

Constantin, 22 mois. L'expérimentateur montre à l'enfant une photo où il porte une tétine.

*CHI : ateti [: tétine]. (mot 1)

*OBS : c'est la tétine de Constantin? (modèle mot 1 + mot 2)

*CHI : Cotintá. [: Constantin]. (mot 2)

*OBS : oui # c'est la tétine de Constantin (modèle mot 1 + mot 2).

Et (2) les SEMeS qui sollicitent un deuxième énoncé par une question :

Exemple n.6. SEMeS inter-tours générées par la conversation

Relation sémantique à travailler : personne + locatif

Paul, 18 mois. L'expérimentateur montre une photo où l'on voit le frère de l'enfant (Louis).

*OBS : qui c'est ?

*CHI : Louis. (mot 1)

*OBS : où est- il ? (solicitation mot 2)

*CHI : là. (mot 2)

*OBS : il est là Louis. (Confirmation mot 1 + mot 2).

Une fois que l'enfant a produit le deuxième mot de la série, l'expérimentateur confirme à la suite la série complète, afin de montrer à l'enfant un modèle de concaténation de deux mots portant sur un même événement-situation.

En ce qui concerne les enfants du groupe contrôle, le matériel et les tâches à réaliser seront les mêmes que ceux du groupe expérimental. Toutefois, les interventions de l'expérimentateur maintiendront le rythme de la communication, sans inciter l'enfant à produire plus d'un mot.

6. Plan expérimental

L'objectif général de cette expérience est d'observer dans le groupe expérimental l'influence du facteur élicitation d'un deuxième élément (entraînement en SEMeS) par la manipulation des mécanismes conversationnels (variable indépendante) sur l'accroissement de la production (variable dépendante) : (1) total de SEMeS et des énoncés à plusieurs mots, ainsi que (2) le total de SEMeS et des productions à deux mots indépendantes de la conversation et du discours. Ces différences s'analysent lors de la comparaison des changements au sein de deux observations naturelles de la conversation enfant-adulte, suite à la présentation (en élicitant ou non les productions des enfants) dans les deux groupes (expérimental et contrôle) d'images construites avec les mêmes critères en fonction de relations sémantiques de leurs éléments.

Pour chaque groupe, deux mesures pre-test sont appliquées : (1) une mesure du vocabulaire cumulatif (CDI), et (2) une mesure quantitative et qualitative des niveaux de dépendance de la conversation et du discours des SEMeS. Ces analyses ont montré que les deux groupes étaient homogènes au départ.

7. Résultats

À partir des données codifiées, nous avons analysé les différences entre la première interaction naturelle et l'interaction naturelle après la présentation d'images, dans les productions de chacun des deux groupes (expérimental et contrôle) à propos des deux variables concernées : (1) production de SEMeS et énoncés à deux mots et (2) production de SEMeS indépendantes de la conversation et du discours.

La différence entre les deux interactions naturelles est le changement. Nous analysons statistiquement le changement par groupe pour comparer ensuite les différences entre groupes à propos de ce changement.

Pour ces analyses, nous nous sommes servi du test de Wilcoxon, étant donné la taille de l'échantillon ($n=10$).

Les résultats ont montré que le test de Wilcoxon pour la différence dans la production des SEMeS et énoncés à deux mots, entre l'interaction naturelle avant les séances et l'interaction après les séances, est significatif dans le groupe expérimental ($T(5)=15$, $P<0.05$, contraste unilatéral), tandis que pour le groupe contrôle, le changement n'est pas statistiquement significatif ($T(4)=12$, $P>0.05$, contraste unilatéral).

En ce qui concerne la production de SEMeS indépendantes de la conversation, du discours et les énoncés à deux mots indépendants du discours, le test de Wilcoxon révèle que le changement dans la production de ces structures est significatif dans le groupe expérimental

($T(5)=16$, $P<0.05$, contraste unilatéral). Au contraire, ce changement n'est pas significatif dans le groupe contrôle ($T(3)=4$, $P>0.05$, contraste unilatéral).

Nous présentons à présent les indices statistiques descriptifs qui illustrent les résultats qualitatifs (cf. Tableau 2).

INTERACTION 2- INTERACTION 1 (= CHANGEMENT)	GROUPE EXPÉRIMENTAL	GROUPE CONTRÔLE
CHANGEMENT DANS LA PRODUCTION TOTAL DES SEMES ET ÉNONCÉS À DEUX MOTS	MOYENNE : 10 MÉDIANE : 9 MAXIMUM : 23 MINIMUM : 1 ÉCART-TYPE : 8,6	MOYENNE : 4,2 MÉDIANE : 1 MAXIMUM : 12 MINIMUM : -2 ÉCART-TYPE : 6,3
CHANGEMENT DANS LA PRODUCTION TOTAL DE SEMES INDEPENDANTES DE LA CONVERSATION ET DU DISCOURS	MOYENNE : 3,2 MÉDIANE : 3 MAXIMUM : 5 MINIMUM : 1 ÉCART-TYPE : 1,5	MOYENNE : 0,6 MÉDIANE : 0 MAXIMUM : 4 MINIMUM : -2 ÉCART-TYPE : 2,2

Tableau 2 - Statistiques descriptives

8. Interprétation et discussion

Les résultats de l'expérience, comme le révèlent les analyses statistiques qui distinguent le groupe expérimental du groupe contrôle sont cohérents avec nos hypothèses de départ. Ils montrent que faciliter la participation active de l'enfant dans des séquences conversationnelles soutenues par l'adulte (SEMeS générées par la conversation) peut avoir un effet positif sur l'accroissement de la production spontanée de ce type de séquences entraînées et des productions à deux mots. On observe cet effet spécialement sur les séquences où l'apport de l'enfant à la conversation est plus autonome (SEMeS et énoncés à deux mots indépendants de la conversation et du discours). Dans ces dernières, l'enfant s'approche de plus en plus d'une production généralisée d'énoncés à deux mots. Ces vérifications empiriques sont cohérentes avec nos suppositions théoriques de départ. L'enfant apprend de l'information qu'il extrait de l'environnement (dans ce cas la stimulation en SEMeS générées par la conversation) et ses propres caractéristiques et capacités de traitement de l'information, infèrent de nouvelles régularités. Cela explique l'accroissement plus élevé des SEMeS indépendantes de la conversation et du discours dans le groupe expérimental, dont la production n'a pas été facilitée directement et qui est l'aboutissement d'une autonomie progressive de l'enfant au sein de la conversation.

Les sujets du groupe expérimental présentent une variabilité importante dans la manière dont ils intègrent cette facilitation. Différents types de séquences et productions transitionnelles sont prédominantes pour chaque enfant.

Références bibliographiques

BATES, Elizabeth & MACWHINNEY, Brian (1987). Competition, variation and language learning, in *Mecanisms of language acquisition*, MacWhinney, Brian (Ed). Hillsdale, N.J. : LEA.

BLOOM, Louis (1973). *One word at a time*. The Hague : Mouton and Co.

BRAINE, Martin (1976). Children's first words combinations. *Monographs of the Society for the Research in Child Development*, 41 (1, Serial No.164).

- BROWN, Roger (1973). *A First Language. The early Stages*. Cambridge : Harvard University Press.
- DORE , John (1985). Holophrases revisited : their logical development from dialog, in *Children's single-word speech*, Barret, Béatrice (Ed). Chichester : Wiley & Sons.
- ELMAN, Jeff L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48, 71-99.
- ELMAN, Jeff L. (1999). The emergence of language : A conspiracy theory, in *The Emergence of Language*, MacWhinney, Brian (Ed). New York : Lawrence Erlbaum Associates.
- ELMAN, Jeff L., BATES, Elizabeth A., JOHNSON, Mark H., KARMILOFF-SMITH, Annette, PARISI, Domenico & PLUNKETT, Kim (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA : MIT Press.
- GREENFIELD, Patricia M. & SMITH, Joshua H. (1976). *The structure of communication in early language development*. New York : Academic Press.
- KARMILOFF-SMITH, Annette (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA : MIT Press/Bradford Books.
- KERN, Sophie (1999). *Inventaire Français du Développement Communicatif chez le nourrisson*. Lyon : Laboratoire Dynamique du Langage. Institut des Sciences de l'Homme.
- LIEVEN, Elena, PINE, John & BALDWIN, James. (1997). Lexically-based learning and early grammatical development, *Journal of Child Language*, 24, 187-219.
- MAC WHINNEY, Brian (1999). *The Emergence of Language*. New York : Lawrence Erlbaum Associates.
- MACWHINNEY, Brian (2000) *The CHILDES Project: Tools for Analyzing Talk. Volume 1: Transcription format and programs*. Volume 2: The Database. Mahwah, New York : Lawrence Erlbaum Associates.
- OCHS, Eleanor, SCHIEFFELIN, Bambi B. & PLATT, Martha (1979). Propositions across utterances, in *Developmental pragmatics*, Ochs, Eleanor & Schieffelin, Bambi B. (Eds). New York : Academic Press.
- SACKS, Harvey, SCHEGLOFF, Emanuel & JEFFERSON, Gail (1974). A simplest systematics for the organisation of turn taking for conversation, *Language*, 50, 696-735.
- SCOLLON, Ron (1979). A early stage: An unzipped condensation of a dissertation on child language, in *Developmental pragmatics*, Ochs, Eleanor & Schieffelin, Bambi B. (Eds). New York : Academic Press, 215-227.
- SCHWARTZ, Richard G., CHAPMAN, Kathy, PRELOCK, Patricia A., TERRELL, Brenda Y. & ROWAN, Lynne E. (1985). Facilitation of early syntax through discourse structure, *Journal of Child Language*, 19, 491-557.
- SNOW, Catherine (1977). The development of conversation between mothers and babies, *Journal of Child Language*, 4, 1-22.
- VENEZIANO, Edy (1989). Les protoconversations et les débuts du langage : de l'alternance des tours de rôle aux significations partagées, *Revue Internationale de Psychologie Sociale*, 2, (1), 23-36.
- VENEZIANO, Edy, SINCLAIR, Hermine & BERTHOUD, Joanna (1990). From one word to two words: repetition patterns on the way to structured speech, *Journal of Child Language*, 17, 633-650.
- VENEZIANO, Edy (1999). Early lexical, morphological and syntactic development in French: Some complex relations, *The international Journal of Bilingualism*, 3, 183-217.
- VIHMAN, Marylin & MCCUNE, Lauren (1994). When a word is a word?, *Journal of Child Language*, 21, 517-542.

SYNTAXE, ACQUISITION ET DIALECTOLOGIE : LA PLURIDISCIPLINARITÉ APPLIQUÉE À L'ÉTUDE DU SUJET NUL ENFANTIN

Katérina PALASIS-JOURDAN

katerina.jourdan@wanadoo.fr

CNRS UMR 6039 "Bases, Corpus, Langage", Université de Nice-Sophia Antipolis
98 bd Edouard Herriot, BP 3209, 06204 Nice Cedex 3

Abstract : A new hypothesis on the child null subject stage is proposed thanks to the association of generative syntax and dialectology. A close analysis of the verbal and pronominal systems of 17 French children around 2 to 6 years old seems to show that, as long as their systems are reduced, they are comparable to those that authorize null subjects (e.g. Italian).

1. La problématique

1.1. Introduction

Le phénomène du sujet nul enfantin est une étape reconnue dans l'acquisition du français. Ce stade a d'ailleurs été déjà assez largement étudié par les générativistes, notamment dans le cadre théorique des Principes et des Paramètres, programme développé à partir de la fin des années 1970 (Chomsky, 1981). En nous plaçant de ce point de vue générativiste, nous proposons aujourd'hui de revenir sur ce phénomène du sujet nul enfantin grâce à l'introduction dans ce cadre théorique d'une approche inhabituelle – et à notre sens complémentaire – qui est celle de la dialectologie. En effet, les théories élaborées par les générativistes sur l'acquisition de la syntaxe se fondent classiquement sur deux éléments. Le premier est la compétence ou la performance linguistique d'un locuteur « idéalisé ». Leurs hypothèses ne proviennent donc pas d'analyses de données empiriques nombreuses. Le second élément est le fait que toutes les hypothèses générativistes sur le sujet nul enfantin postulent une déficience de la compétence ou de la performance de l'enfant par rapport à celle de l'adulte. Nous introduisons donc le terme de dialectologie dans un double but méthodologique. Tout d'abord, cette discipline linguistique s'intéresse à l'étude de systèmes linguistiques *oraux*, par le biais de *corpus* collectés sur le *terrain* (Dalbera, 2002). Certes, rassembler des données orales n'est ni nouveau ni réservé à la dialectologie – les psycholinguistes par exemple partagent cette approche avec les dialectologues. Cependant, les méthodes et les points de vue divergent. Ainsi, le second but que nous poursuivons en adoptant ce point de vue est d'analyser les productions enfantines en tant que *système* linguistique *à part entière* et de comparer ce système à celui de l'adulte afin de dégager les différences qui expliqueraient l'existence du phénomène du sujet nul dans le premier système et pas dans le second¹. Nous n'adhérons donc pas totalement au point de vue générativiste dans la mesure où nous attribuons une origine intra-systémique au phénomène et que nous ne postulons donc pas une déficience enfantine quelconque.

¹ Nous passons ici sur le fait que le français possède tout de même des verbes défectifs comme falloir où le sujet peut être omis à l'oral (Haegeman, 1997), sur l'existence des impératifs qui ne présentent pas non plus de sujets, et sur la réduction phonologique de nombreux pronoms dans l'input de l'enfant. Le français adulte n'est donc pas complètement exempt de sujets nuls (Kayne, 2000 et Oliiviéri, 2004 sur la variation micro-paramétrique).

Mais avant d’appliquer cette association entre le cadre théorique générativiste et le travail de terrain du dialectologue, nous nous proposons tout d’abord de rappeler précisément ce qu’est ce phénomène du sujet nul enfantin et quelles hypothèses existent déjà sur ce thème dans le cadre de la grammaire générative. Dans un second temps, après avoir identifié ce que nous considérons être les limites de ces hypothèses, nous exposerons ainsi notre proposition nouvelle sur cette question.

1.2. Qu’est-ce que le phénomène du sujet nul enfantin ?

L’étude du phénomène du sujet nul enfantin repose sur l’existence du Paramètre du Sujet Nul. Ce paramètre établit une dichotomie entre, d’une part, les langues « à sujet nul » (comme l’italien ou l’espagnol) et, d’autre part, les langues « à sujet non-nul » (comme le français ou l’anglais). Les critères généralement admis par les générativistes pour établir cette dichotomie sont au nombre de quatre ; ils ont été formalisés notamment par Rizzi (1997). Pour ne citer que les deux critères les plus étroitement liés au phénomène du sujet nul enfantin, nous pouvons mentionner :

- La présence optionnelle *vs* obligatoire d’un sujet préverbal pronominal. Rizzi oppose ainsi l’italien *parla* au français *il parle*.
- La richesse *vs* la pauvreté morphologique du verbe. L’auteur oppose ainsi le verbe italien *parlare* au présent de l’indicatif (*parlo, parli, parla, parliamo, parlate, parlano*) qui comporte six formes phonétiquement différentes sur les six du paradigme, au verbe français *parler* qui, lui, ne comporte que trois formes phonétiquement différentes (*parl, parlons, parlez*).

Avec cette base théorique comme point de référence, il a néanmoins été observé que les enfants qui sont en train d’acquérir une langue « à sujet non-nul » (comme le français) omettent quand même et systématiquement jusqu’à l’âge approximatif de trois ans certains de leurs sujets. Ce sont ces omissions qui sont appelées « phénomène ou stade du sujet nul enfantin ». Pour ce qui est du français, pour lequel nous proposons quelques exemples extraits de notre corpus dans le Tableau 1, ce phénomène semble survenir dans environ un quart des phrases enfantines (Rasetti, 1996 : 26,2 % ; Palasis-Jourdan, 2005 : 22,4 %).

Phrases enfantines	Équivalents adultes	Prénoms et âges des enfants ²
* <i>est vert.</i>	« c’ est vert »	Raphaël, 2;6
* <i>oui veux Toupie.</i>	« oui je veux Toupie »	Matteo, 2;11
* <i>fais la moto.</i>	« je fais la moto » (pâte à modeler)	Thibault, 2;9
* <i>a lion.</i>	« il y a un lion »	Thibault, 2;9
* <i>sais pas.</i>	« je (ne) sais pas »	Maxence, 2;9

Tableau 1 - Quelques exemples d’énoncés enfantins à sujets nuls (Palasis-Jourdan, 2005)

Il est important de noter que les phrases ci-dessus ont été produites par des enfants francophones qui ne présentent aucune pathologie ou trouble particulier du langage et, pourtant, nous constatons qu’il manque « quelque chose » dans ces phrases. Pour utiliser la terminologie usuelle et parce que ce n’est pas le propos de discuter ce point ici, nous appellerons cet élément omis « sujet », à la suite des appellations existantes que sont « Paramètre du Sujet Nul » et « phénomène du sujet nul enfantin ». Cependant, ces appellations peuvent être discutées dans la mesure notamment où cette notion de « sujet » est complexe et qu’elle ne précise pas la nature de l’élément mais uniquement sa fonction grammaticale. La question de la nature du sujet suscite d’ailleurs différents débats. Deux illustrations en sont notamment la distinction qui est proposée entre l’oral et l’écrit (Blanche-Benveniste, 1994, 2003) et la discussion qui existe sur le statut de ce qui est communément appelé « pronom personnel sujet » : s’agit-il d’un argument syntaxique à part entière

² Les âges sont donnés en années;mois.

(Chomsky, 1981, Côté, 2001, De Cat, 2005) ou d'un affixe flexionnel du verbe (Pierce, 1992, Kaiser, 1994, Creissels, 1995) ? Mais pour l'heure, revenons aux hypothèses générativistes sur le phénomène du sujet nul enfantin.

1.3. Les hypothèses générativistes

Ces hypothèses invoquent généralement l'un ou l'autre des éléments de la dichotomie chomskyenne traditionnelle : la compétence ou la performance linguistique. Plus précisément, elles attribuent les omissions enfantines à un déficit soit de la compétence infantine soit de la performance infantine comparée à celle de l'adulte. De nombreux travaux ont été publiés dans ce domaine. En ce qui concerne les hypothèses proposant un déficit de la compétence, nous pouvons mentionner les travaux fondateurs de Hyams (1986) ou encore ceux de Rizzi (2002) qui postulent tous les deux un mauvais paramétrage initial. Pour ce qui est du déficit de la performance, les sujets nuls enfantins sont attribués à des capacités limitées des enfants au niveau de la mémoire, de la computation et/ou de la production. Jakubowicz & Rigaut (1997), par exemple, évoquent une réduction phonologique plus ou moins importante des pronoms sujets.

Cependant, il nous semble que ces différentes hypothèses peuvent conduire à une même interrogation, voire à une même objection. En effet, nous avons observé que le phénomène du sujet nul enfantin ne concerne pas la totalité – ni même la majorité – des productions enfantines. La question que nous posons est donc la suivante : si l'on attribue les phrases sans sujet à un déficit de la compétence ou de la performance de l'enfant comparée à celle de l'adulte, comment explique-t-on, dans le cadre de la même hypothèse, que l'enfant soit non seulement capable de produire également des phrases avec sujet mais que, de plus, ces phrases avec sujet représentent la majorité de sa production ? Les hypothèses générativistes proposent ainsi bien sûr chacune une explication au phénomène du sujet nul mais il nous semble qu'aucune de ces solutions ne rend compte de la présence concomitante des deux types de structures chez le jeune enfant et de la proportion des sujets non-réalisés.

2. Notre proposition

Face à ce que nous pensons être les limites de ces hypothèses, nous proposons donc d'ouvrir l'horizon théorique de la grammaire générative à celui de la dialectologie. Nous sortons ainsi de la dichotomie traditionnelle compétence/performance du locuteur idéalisé de la grammaire générative en n'adhérant pas notamment au postulat qui consiste à poser que le système linguistique du jeune enfant présente une déficience par rapport au système adulte. Nous nous intéressons plutôt à un troisième facteur : celui de la variation qui existe entre le système linguistique du jeune enfant et celui de l'adulte. Et pour identifier ces différences, il faut bien sûr travailler à partir de données adultes et enfantines. En ce qui concerne le système linguistique de l'adulte, nous avons basé nos observations sur les travaux de Blanche-Benveniste (1994, 2003) sur le français oral. Pour ce qui est des données enfantines, nous avons constitué un corpus de productions spontanées, grâce à cinq séances d'enregistrements effectuées entre octobre 2003 et janvier 2004 dans une classe de première année de maternelle. Les productions finalement analysées proviennent ainsi de dix-sept enfants qui avaient entre 2;3 (âge du plus jeune enfant lors de la première séance) et 3;1 (âge de l'enfant le plus âgé lors de la dernière séance). Dans le cadre de cette étude, nous n'avons analysé que les phrases comportant un verbe, soit un total de 1072 phrases verbales.

Notre proposition sur le sujet nul enfantin s'articule alors autour de deux points principaux. Dans un premier temps, nous postulons que le jeune enfant, au cours de ses premiers mois, attribue la valeur correcte au Paramètre du Sujet Nul. Les enfants francophones fixent ainsi correctement la valeur du paramètre sur « langue à sujet non-nul » grâce à ce qu'ils entendent dans leur environnement linguistique. Ce postulat permet de rendre compte des

trois-quarts des phrases produites par les jeunes enfants, à savoir celles qui sont grammaticalement correctes en ce qui concerne la présence du sujet. Ce postulat d’une fixation correcte et précoce – c’est-à-dire avant le moment où le jeune enfant entre dans la phase de combinaisons de deux mots (vers vingt mois) – va donc à l’encontre des hypothèses de Hyams (1986) ou Rizzi (2002), qui eux proposent une valeur initiale erronée. Nous n’adoptons pas ce point de vue car une telle hypothèse ne nous semble pas conforme aux principes récurrents d’économie dans l’acquisition du langage. Il nous semble en effet beaucoup plus simple de supposer une fixation unique et correcte des paramètres plutôt que de poser que les paramètres sont initialement fixés de façon erronée et que ces valeurs doivent être corrigées par la suite. De plus, cette hypothèse nécessite également d’expliquer pour quelle(s) raison(s) le jeune enfant débiterait son acquisition muni d’éléments erronés. Par contre, notre point de vue est similaire à celui de Wexler (1998) par exemple qui postule que les paramètres de base (tels que ceux qui régissent l’ordre des éléments, le mouvement du verbe fléchi ou le sujet nul) sont fixés tôt au cours de l’acquisition. Notre deuxième point principal consiste alors en l’explication des productions restantes, celles qui ne sont pas conformes à la valeur du Paramètre du Sujet Nul. Nous nous tournons donc vers le système linguistique de l’enfant, grâce aux données recueillies sur le terrain, avec pour objectifs d’y trouver une explication :

- qui rende compte des sujets nuls enfantins,
- qui soit compatible avec la présence concomitante de sujets réalisés,
- qui apporte une explication à la proportion des sujets nuls *vs* des sujets réalisés.

Nous proposons ainsi une analyse en trois étapes : étude du système verbal enfantin, étude du système pronominal enfantin, croisement des observations recueillies au cours des deux premières étapes.

2.1. Étude du système verbal enfantin

Le Tableau 2 ci-dessous présente une synthèse de l’analyse du système verbal des enfants enregistrés. Il donne la liste, par ordre décroissant de fréquence, des verbes majoritairement présents dans le corpus. Dans un souci de représentativité, nous avons posé des filtres sur cette sélection de verbes. Cette liste exclut ainsi les verbes qui n’ont été produits que par un seul enfant du groupe et, de fait, les verbes utilisés une seule fois dans le corpus. Cette liste exclut également les verbes produits à l’impératif dans la mesure où ce mode ne nécessite jamais la présence d’un sujet.

Verbes	Effectifs	%	Groupes	Effectifs	%
<i>être</i>	286	48,7	3 ^{ème} et <i>être/avoir</i>	557	94,9
<i>vouloir</i>	101	17,2			
<i>faire</i>	95	16,2			
<i>avoir</i>	41	7,0			
<i>savoir</i>	21	3,6			
<i>voir</i>	13	2,2			
<i>manger</i>	9	1,5	1 ^{er}	30	5,1
<i>rouler</i>	6	1,0			
<i>jouer</i>	5	0,9			
<i>tomber</i>	4	0,7			
<i>cacher</i>	3	0,5			
<i>marcher</i>	3	0,5			
Corpus filtré	587	100		587	100

Tableau 2 - Le système verbal enfantin (verbes principaux, par ordre décroissant de fréquence)

Un premier phénomène à relever est le fait qu’aucune différence n’a été remarquée entre les verbes présents dans les phrases avec sujet et ceux présents dans les phrases sans sujet. La piste du type de verbe pour expliquer le phénomène du sujet nul enfantin n’a donc pas été retenue. Par contre, ce Tableau 2 met en avant un autre phénomène qui nous paraît

intéressant. Il s'agit du fait que l'usage enfantin est presque entièrement composé de verbes du troisième groupe. Ce phénomène n'est, en lui-même, pas complètement surprenant dans la mesure où certains de ces verbes sont extrêmement fréquents. Ce que nous souhaitons mettre en avant ici, en rapport avec ce groupe de verbes, est le fait que ces verbes du troisième groupe présentent dans leurs paradigmes du présent de l'indicatif un nombre de formes phonétiquement différentes plus élevé que les verbes du premier groupe – groupe utilisé par Rizzi (1997), entre autres, pour poser la dichotomie sur la richesse ou la pauvreté de la morphologie verbale.

Le Tableau 3 détaille ces paradigmes verbaux. Nous constatons ainsi que *être* et *avoir* présentent cinq formes différentes sur les six du paradigme, que *vouloir*, *faire* et *savoir* en ont quatre sur six alors que *voir* et les verbes du premier groupe n'en présentent que trois.

Personnes	<i>être</i>	<i>vouloir</i>	<i>faire</i>	<i>avoir</i>	<i>savoir</i>	<i>voir</i>	1 ^{er} groupe
1sg	sWi	vĖ	fe/fE	e/E	se/sE	vwa	m\$AZ
2sg	e/E	vĖ	fe/fE	a	se/sE	vwa	m\$AZ
3sg	e/E	vĖ	fe/fE	a	se/sE	vwa	m\$AZ
1pl	sOm	vul\$O	fêz\$O	av\$O	sav\$O	vwaj\$O	m\$AZ\$O
2pl	Et	vule	fEt	ave	save	vwaje	m\$AZe
3pl	s\$O	vœl	f\$O	\$O	sav	vwa	m\$AZ
% corpus	48,7	17,2	16,2	7,0	3,6	2,2	5,1
Formes #	5	4	4	5	4	3	3

Tableau 3 - Les paradigmes complets des verbes présents dans le corpus enfantin (API)

Cette première analyse permet donc de constater que le système verbal enfantin porte beaucoup moins bien l'étiquette de « pauvreté morphologique » que son équivalent adulte, pour deux raisons :

- les verbes du premier groupe sont quasiment absents de ce système (5,1 %),
- la majorité des verbes utilisés présente quatre ou cinq formes différentes (92,7 %).

Ces caractéristiques nous rapprochent alors beaucoup de la morphologie verbale riche des langues à sujet nul qui, elle, présente cinq ou six formes différentes.

2.2. Étude du système pronominal enfantin

Nous nous tournons vers la deuxième étape de notre analyse avec l'étude du système pronominal enfantin. En effet, une question intéressante est de se demander maintenant si les enfants utilisent toutes les formes des paradigmes détaillés dans le Tableau 3 ou seulement quelques-unes d'entre elles. Le Tableau 4 présente ainsi la synthèse de notre analyse du système pronominal enfantin. Les effectifs mentionnés indiquent le nombre de phrases comportant ou susceptibles de comporter chaque pronom de la liste (c'est-à-dire que, lorsque le pronom a été omis par l'enfant, nous l'avons reconstitué d'après le contexte). Ces données ont également été filtrées : toutes les formes phonétiquement ambiguës ont été écartées, ainsi que les formes à l'impératif.

Pronoms	Effectifs	%	Personnes	%	%
<i>je</i>	269	35	1sg	35	98,8
<i>tu</i>	22	2,9	2sg	2,9	
<i>il</i>	180	23,4	3sg	60,9	
<i>il explétif</i>	17	2,2			
<i>elle</i>	33	4,3			
<i>c'</i>	180	23,4			
<i>ça</i>	19	2,5			
<i>on</i>	39	5,1			
<i>nous</i>	0	0	1pl	0	1,2
<i>vous</i>	0	0	2pl	0	
<i>ils</i>	6	0,7	3pl	1,2	
<i>elles</i>	4	0,5			
Corpus filtré	769	100		100	100

Tableau 4 - Répartition des pronoms et des personnes utilisés par les enfants

Ce tableau permet de remarquer que ce système pronominal est composé, dans sa quasi-totalité, de deux formes seulement :

- la forme de la première personne du singulier (avec le pronom *je*) dans 35 % des phrases,
- la forme de la troisième personne du singulier (avec différents pronoms) dans 60,9 % des cas.

Ce système pronominal peut donc être décrit comme :

- étant très majoritairement singulier (à 98,8 %),
- ne comportant qu’une seule opposition entre deux formes à l’intérieur de ce système du singulier : la première personne (que nous appelons, à la suite de Benveniste (1966), [+Locuteur]) et la troisième personne ([–Locuteur]).

De nouveau donc, comme pour le système verbal, nous remarquons que le système pronominal des enfants présente des limitations importantes par rapport à toutes les possibilités du paradigme adulte.

2.3. Croisement des observations

Le Tableau 5 ci-dessous présente le résultat du croisement entre les informations obtenues dans les Tableaux 3 et 4, c’est-à-dire que nous retrouvons les verbes majoritairement utilisés par les enfants mais avec les paradigmes réduits aux deux personnes principales de leur système pronominal.

Personnes	<i>être</i>	<i>vouloir</i>	<i>faire</i>	<i>avoir</i>	<i>savoir</i>	<i>voir</i>	1 ^{er} groupe
1sg	sWi	vE	fe/fE	e/E	se/sE	vwa	m§AZ
3sg	e/E	vE	fe/fE	a	se/sE	vwa	m§AZ
% corpus	48,7	17,2	16,2	7,0	3,6	2,2	5,1
Formes ≠	2	1	1	2	1	1	1

Tableau 5 - Données croisées (verbes utilisés mais réduits aux deux personnes du système, API)

Nous observons alors que, dans plus de la moitié des phrases enfantines (avec les verbes *être* et *avoir*), nous retrouvons une morphologie verbale riche puisque l’opposition entre les deux formes *suis* et *est*, *ai* et *a* est totalement distinctive, comme dans une langue comme l’italien qui présente, pour les mêmes verbes, les formes *sono* et *è*, *ho* et *ha*. Et donc, comme en italien par exemple, la forme verbale se suffit à elle-même, c’est-à-dire qu’elle ne nécessite pas la présence obligatoire d’un sujet pour que la référence à l’argument principal soit totalement compréhensible. Une certaine absence du sujet dans ce système enfantin nous apparaît alors comme compatible avec la fonctionnalité communicative de l’énoncé.

3. Conclusion

L'utilisation du cadre théorique de la grammaire générative (avec la théorie des Principes et des Paramètres) élargi à la méthodologie de la dialectologie (étude et comparaison de systèmes linguistiques en tant que tels) nous permet ainsi d'émettre une hypothèse nouvelle sur le sujet nul enfantin. En effet, grâce à notre étude, nous constatons que le jeune enfant se trouve dans une situation linguistique que nous qualifions de duale. D'un côté, le système linguistique de l'adulte francophone est un système « à sujet non-nul ». Au cours de ses premiers mois, le jeune enfant fixe donc la valeur du Paramètre du Sujet Nul de façon correcte grâce aux informations qu'il reçoit de son entourage (input). Ce paramétrage correct lui permet ainsi de produire environ 75 % de phrases qui se conforment au modèle adulte en ce qui concerne la présence du sujet. Cependant, de l'autre côté, notre analyse met en relief le fait que le système linguistique du jeune enfant (output) présente des formes verbales distinctives (*suis/est* et *ai/a*), ce qui est en principe l'une des caractéristiques de l'autre groupe de langues, celles « à sujet nul ». C'est cette différence entre l'input et l'output que nous identifions comme étant le déclencheur de la présence d'environ 25 % de sujets nuls chez le jeune enfant.

Ainsi, si nous reprenons nos trois objectifs mentionnés dans la première partie du paragraphe 2, notre hypothèse :

- rend compte des sujets nuls enfantins (input vs output),
- est compatible avec la présence des sujets réalisés (valeur correcte du paramètre),
- rend compte de la proportion modérée de sujets nuls dans la mesure où le déclencheur est un fait relativement discret dans le système puisque seulement deux verbes présentent une morphologie verbale totalement distincte chez l'enfant à ce stade ; son impact est donc quantitativement limité.

Par ailleurs, nous proposons également que cette hypothèse pourrait être étendue à d'autres langues qui présentent ce phénomène du sujet nul enfantin, s'il était démontré que les jeunes enfants anglophones ou germanophones par exemple utilisent essentiellement, à ce stade, des verbes irréguliers aux personnes 1 et 3 (anglais : *am* et *is*, *have* et *has* ; allemand : *bin* et *ist*, *habe* et *hat*). Notre hypothèse présente également l'avantage de pouvoir expliquer comment l'enfant quitte ce stade du sujet nul enfantin vers 3;0. En effet, dans ce cadre, la disparition de ce phénomène peut être attribuée à l'acquisition de nouveaux pronoms et de nouveaux verbes, et notamment des verbes du premier groupe qui ne présentent plus cette richesse morphologique. Le système de l'enfant s'étoffe ainsi petit à petit pour se rapprocher progressivement du système de l'adulte. La différence entre les deux systèmes linguistiques – identifiée comme étant le déclencheur du phénomène – s'atténue alors et le phénomène disparaît.

Références

- BENVENISTE, Émile (1966). *Problèmes de linguistique générale, 1*, Paris : Gallimard.
- BLANCHE-BENVENISTE, Claire (1994). Quelques caractéristiques grammaticales des "sujets" employés dans le français parlé des conversations, in *Subjecthood and Subjectivity. Proceedings of the Colloquium "The Status of the Subject in Linguistic Theory"*, London, 19-20 March 1993, Yaguello, Marina (Ed), Gap, Paris & Londres : Éditions Ophrys & Institut français du Royaume-Uni, 77-107.
- BLANCHE-BENVENISTE, Claire (2003). Les formes grammaticales de réalisation des sujets et leur inégale représentation en français contemporain, in *Le sujet*, Merle, Jean-Marie (Ed). Gap & Paris : Éditions Ophrys, 73-90.
- CHOMSKY, Noam (1981). *Lectures on Government and Binding*. Dordrecht : Foris.
- CÔTÉ, Marie-Hélène (2001). On the Status of Subject Clitics in Child French, in *Proceedings of the 8th Conference of the IASCL*, Almgren, Margareta, Barrena, Andoni, Ezeizabarrena,

- Maria-José, Idiazabal, Itziar & MacWhinney, Brian (Eds). Somerville, MA : Cascadilla Press, 1314-1330.
- CREISSELS, Denis (1995). *Éléments de syntaxe générale*. Paris : Presses Universitaires de France.
- DALBERA, Jean-Philippe (2002). Le corpus entre données, analyse et théorie, *Corpus*, 1 : 89-104.
- DE CAT, Cécile (2005). French Subject Clitics are not Agreement Markers, *Lingua*, 115, 9 : 1195-1219.
- HAEGEMAN, Liliane (1997). Adult Null Subjects in Non Pro-drop Languages, in *The Acquisition of Syntax. Studies in Comparative Developmental Linguistics*, Friedemann, Marc-Ariel & Rizzi, Luigi (Eds). London : Longman (2000), 129-169.
- HYAMS, Nina M. (1986). *Language Acquisition and the Theory of Parameters*. Dordrecht : D. Reidel Publishing Company.
- JAKUBOWICZ, Célia & RIGAUT, Catherine (1997). L'acquisition des clitiques nominatifs en français, in *Les pronoms : morphologie, syntaxe et typologie*, Zribi-Hertz, Anne (Ed). Saint-Denis : Presses Universitaires de Vincennes, 57-99.
- KAISER, Georg A. (1994). More about INFL-lection and Agreement. The Acquisition of Clitic Pronouns in French, in *Bilingual First Language Acquisition: French and German Grammatical Development*, Meisel, Jürgen M. (Ed). Amsterdam & Philadelphia : John Benjamins, 133-159.
- KAYNE, Richard S. (2000). Microparametric Syntax. Some Introductory Remarks, in *Parameters and Universals*, Kayne, Richard S. (Ed). Oxford : Oxford University Press, 3-9.
- OLIVIÉRI, Michèle (2004). Y a-t-il des frontières dialectales en syntaxe ?, *Études corses*, 59 : 77-93.
- PALASIS-JOURDAN, Katérina (2005). *Problèmes d'acquisition et le Paramètre du Sujet Nul*. Mémoire de Master, Université de Nice-Sophia Antipolis.
- PIERCE, Amy (1992). *Language Acquisition and Syntactic Theory: A Comparative Analysis of French and English Child Grammars*. Dordrecht : Kluwer Academic Publishers.
- RASETTI, Lucienne (1996). Null Subjects and Root Infinitives in the Child Grammar of French, *Geneva Generative Papers*, 4, 2 : 120-132.
- RIZZI, Luigi (1997). A Parametric Approach to Comparative Syntax: Properties of the Pronominal System, in *The New Comparative Syntax*, Haegeman, Liliane (Ed). Harlow : Addison Wesley Longman, 268-285.
- RIZZI, Luigi (2002). *On the Grammatical Basis of Language Development: A Case Study*. Ms., University of Siena.
- WEXLER, Ken (1998). Very Early Parameter Setting and the Unique Checking Constraint: A New Explanation of the Optional Infinitive Stage, *Lingua*, 106, 1-4 : 23-79.

FORMES PAR DÉFAUT ET FRÉQUENCE DANS L'INPUT EN FLE

Anita THOMAS

anita.thomas@rom.lu.se

Université de Lund – SOL français - Box 201

SE-221 00 Lund

Suède

Abstract : This article examines the production of forms like *je parlE* by Swedish adult learners of French. It highlights the contribution of the Missing Surface Inflection Hypothesis regarding the use of default forms. Based on an analysis of the frequency of surface forms in the input, this study will further show that the variation found in learners' interlanguage does show some systematic aspects.

1. Introduction

Le projet de thèse présenté dans cet article s'inscrit dans la recherche sur l'acquisition du français langue étrangère (FLE), en particulier de la morphologie verbale, en production orale. Les analyses sont basées sur des données d'apprenants adultes suédophones¹. Cette étude est centrée sur les verbes réguliers en *-er* dans les contextes du présent de l'indicatif et de l'infinitif. Je m'intéresse tout particulièrement à la production de formes interlangagières du type *je parlE*² ou *je veux habite* comme dans les exemples ci-dessous :

Ex. (1) : *Forme longue dans un contexte pour une forme du présent :*

*INT: *mais qu'est ce que vous faites alors maintenant que vous êtes retraité?*

*BER1: *euh [/] euh je travaille à la euh jardin .*

Ex. (2) : *Forme du présent dans un contexte pour une forme de l'infinitif*

*INT: *mais vous n'allez pas habiter en France?*

*JUL: *peut être en quatre ans. (...)*

*JUL: *euh nous allons habite en euh Perpignan ...*

De nombreuses recherches ont été consacrées à la production de telles formes, notamment dans le cadre de la grammaire générative. Une telle approche pose la question du statut syntaxique des formes interlangagières et propose que les apprenants adultes débutants ont recours à des formes *par défaut*. L'objectif de mon propre projet est d'approfondir la question de la production de possibles formes par défaut à partir de la question de l'influence de la fréquence des formes dans l'input et de leur distribution, proposant ainsi que la variation des formes interlangagières reflète une certaine systématisme, les formes courtes du présent (*parl-*) étant particulièrement liées aux verbes d'état et les formes longues (*parlE*) aux verbes dynamiques.

Cet article présente la première partie de mon projet de thèse (Thomas, 2006), basé sur l'analyse d'un corpus de conversations. Ces premiers résultats serviront de base à la formulation d'hypothèses plus précises qui seront vérifiées dans des tests expérimentaux.

¹ Le choix d'un corpus de suédophones est lié au fait que mon projet de thèse s'inscrit dans la recherche sur le FLE en Suède.

² Par convention le *-E* correspond à un son final en /e/. Une telle transcription permet de garder l'ambiguïté de la forme visée par l'apprenant, *parler, parlé, parlez, parlais/t/ent* ou, dans le cas d'une difficulté de prononciation, *parle*.

2. Considérations théoriques

2.1. *Apport de la grammaire générative*

La question centrale dans le cadre de la grammaire générative est de savoir si la production de formes interlangagières pourrait être due à une difficulté au niveau de la syntaxe des apprenants de L2/LE ou, en d'autres termes, si les apprenants ont accès ou non aux catégories fonctionnelles et plus particulièrement au trait abstrait [\pm fini]. En effet, dans cette approche, on considère que la production de formes finies est un indice de l'accès à la catégorie fonctionnelle *Inflection*, qui contient les traits syntaxiques du verbe conjugué, notamment le temps et la finitude mais aussi la personne et le nombre (Hawkins, 2004). Le verbe lexical est ainsi soumis à une vérification des traits grammaticaux de la fonction finie visée. Le verbe non-fini par contre reste dans la phrase verbale VP, le point d'insertion lexical. (Prévost & White, 2000).

L'hypothèse de l'absence de fréquence de surface, *Missing Surface Inflection Hypothesis* – MSIH propose que la difficulté pour les apprenants adultes se situe, non pas au niveau de la syntaxe, mais au niveau de la réalisation de la morphologie de surface (Prévost & White, 2000 : 103). Pour vérifier cette hypothèse, Prévost et White ont procédé à un certain nombre de tests dont le placement de la négation. Si les formes du type *parlE* apparaissent plus fréquemment à gauche de la négation *pas* (*je parlE pas*) qu'à droite (*je pas parlE*) on peut conclure que la forme verbale est montée dans *Inflection*, la place hiérarchique de la négation étant au-dessus de la phrase verbale VP et en dessous de la catégorie *Inflection*.

Les données analysées par Prévost et White semblent indiquer que les formes non-finies produites dans des contextes demandant une forme finie (*je parlE*) possèdent toutes les caractéristiques syntaxiques des formes finies (Prévost & White, 2000; Prévost, 2004). La difficulté se situe donc au niveau de la réalisation de la morphologie de surface, la forme non-finie servant de forme *par défaut*.

La production de formes finies dans un contexte demandant une forme non-finie (*je veux habite*) n'a pas été observée dans les données de Prévost et White. Elle l'est par contre dans Prévost (2004), une nouvelle étude transversale faite avec des apprenants anglophones, répartis en quatre niveaux de français. Comme pour le phénomène précédent, il observe une baisse de la sur-utilisation des formes finies avec l'augmentation du niveau de français. Ainsi, à l'instar de la production de formes non-finies dans une position finie (*je parlE*), les formes finies en position non-finie (*je veux parle*) ne sont finies qu'en surface portant toutes les caractéristiques d'une forme non-finie. L'acquisition de la syntaxe et de la morphologie se ferait donc de manière indépendante pour les apprenants adultes, la syntaxe étant acquise dès le début.

Bien que la présente étude ne se situe pas dans le cadre de la grammaire générative, les recherches dans celui-ci, et en particulier celles de Prévost et White, forment une étape antérieure importante. En effet, les résultats de leurs recherches permettent d'exclure le problème du déficit syntaxique et montrent que les apprenants débutants ont recours à des formes par défaut. Cependant la question des causes de la variation des formes dans l'interlangue des apprenants reste intacte. En effet, le fait d'avoir accès aux catégories fonctionnelles ne permet pas aux apprenants de produire la morphologie verbale appropriée. Il est donc nécessaire de poursuivre la recherche sur d'autres influences possibles.

Dans la présente recherche, je suivrai une des suggestions de Hawkins (2004) qui consiste à considérer la question comme pouvant être l'effet d'un phénomène extérieur à la grammaire générative, comme la fréquence dans l'input.

2.2. Influence de la fréquence dans l'input

Les recherches centrées sur les effets de la fréquence de l'input mettent l'accent sur le fonctionnement général de la cognition humaine, faisant de l'apprenant d'une LE/L2 un locuteur comme un autre. Selon ces théories (Bybee, 1991 ; Ellis, 2002), la cognition humaine fonctionnerait de telle sorte que les occurrences trouvées au fil de l'expérience de la langue (*usage-based*) seraient regroupées en catégories par association de traits similaires. La répétition de certaines occurrences aurait en outre l'effet de renforcer la représentation de celles-ci. Finalement, la sensibilité à la fréquence implique une propension au calcul de probabilité d'occurrence.

Cette manière de traiter les informations perçues dans l'input peut s'illustrer par le principe des tendances distributionnelles qui se définit de la manière suivante : si les formes X et Y peuvent apparaître dans les deux contextes A et B, mais que les tendances distributionnelles font percevoir que X apparaît surtout dans A et Y dans B, alors les apprenants restreindront l'utilisation de X au contexte A et de Y au contexte B. (Mitchell & Myles, 1998). Le tableau 1 ci-dessous illustre ce principe selon l'interprétation que j'en fais. Les formes en gras sont les formes les plus fréquentes dans chacun des contextes et ce sont les formes que les apprenants reprennent dans leur interlangue.

	Contexte A	Contexte B
fréquence prototypique	forme X et forme Y	forme X et forme Y
exemples locuteurs natifs	elle habite / elle va habiter	il tombe / il va tomber
exemples interlangue	elle habite / *elle va habite	*il tombE / il va tombE

Tableau 1 - Principe des tendances distributionnelles

Selon ce principe, on peut observer une corrélation entre la distribution des formes et des fonctions dans l'input et la production des apprenants. En d'autres termes, ce principe fait l'hypothèse d'une correspondance entre la (proto)typicité dans l'input et le choix des formes chez les apprenants.

Bybee (1991) propose que les apprenants d'une L2/LE commencent par acquérir les formes les plus fréquentes et sémantiquement les plus fondamentales dans un paradigme, ce qui pour des apprenants du FLE pourrait correspondre aux formes de l'infinitif et du présent singulier.

En effet, dans un énoncé (*je veux voir ce film*), l'infinitif est souvent l'élément central pour la communication d'un contenu, il est porteur de l'information lexicale. En outre, l'utilisation de l'infinitif comme forme de citation le rend extrêmement fréquent.

D'autre part, la forme du présent singulier peut être considérée comme une *forme de base*, servant de base à la construction des autres formes d'un même paradigme. Les formes marquées d'un même paradigme peuvent ainsi être vues comme étant en opposition avec cette forme de base (Bybee, 1985). Pour les verbes en *-er* en français, la forme de base correspond à la forme courte du présent (*parl-*). En suédois par contre, la forme de base serait plutôt la forme de l'infinitif (*prata 'parler'*) que la forme du présent (*pratar 'parle'*).

Le processus de simplification du système morphologique de la langue cible par les apprenants débutants, tel qu'illustré par le principe des tendances distributionnelles, pourrait ainsi aboutir à la répartition de la forme courte du présent et de la forme longue de l'infinitif (et de ses homonymes) en deux catégories, influencé par la fréquence des formes dans l'input. De nombreuses études à partir des catégories sémantiques proposées par Vendler (1957), notamment sur le FLE, ont mis en évidence le lien privilégié entre le présent et les verbes d'état (cf. par exemple Kihlstedt, 1998).

On peut ainsi formuler l'hypothèse que la forme courte du présent est avant tout associée aux verbes d'état et, par opposition, la forme longue en /e/ aux verbes dynamiques dans l'input comme chez les apprenants débutants.

3. Hypothèse

Ainsi, à partir des théories sur l'influence de la fréquence des formes dans l'input et le principe des tendances distributionnelles, je propose l'hypothèse suivante :

Pour les verbes réguliers en *-er*, l'utilisation et la sur-utilisation de la forme courte du présent et de la forme longue se terminant par le son /e/ par les apprenants du FLE, reflète la distribution de ces formes dans l'input. Dans celui-ci, la fréquence de la forme courte du présent (singulier et 3^e du pluriel) d'une part et des formes se terminant en /e/ (infinitif, participe passé, imparfait, 2^e pluriel du présent) d'autre part, se distribuent de telle sorte que la forme courte du présent est la forme la plus fréquente pour les verbes d'état et la forme en /e/ la plus fréquente pour les verbes dynamiques.

4. Corpus

Les données de la production des apprenants proviennent de deux corpus. Tout d'abord d'un corpus pilote que j'ai constitué en 2003-2004, comprenant 20 apprenants adultes suédophones (moyenne d'âge 57 ans) ayant eu entre 24 et 96 heures de cours du soir de français. Près de la moitié des apprenants avaient déjà appris le français à l'école dans leur jeunesse. Tous ont l'anglais comme première langue étrangère. Les enregistrements, comprenant des conversations semi-spontanées, sont d'une durée allant de 10 à 30 minutes et certains apprenants ont été enregistrés plusieurs fois. Au total, le corpus contient ainsi 25 enregistrements. Ce corpus pilote a été complété par les données de deux apprenants du Corpus Lund, Karl (23 ans), 3 enregistrements (sur 5) et Lisa (20 ans), 4 enregistrements, tous d'une durée de 60 minutes environ. Ces enregistrements comprennent avant tout des conversations libres mais aussi un certain nombre de tâches (raconter un conte, décrire un chemin). Karl est un apprenant qui a essentiellement appris le français sur place en France à son travail, alors que Lisa a appris le français à l'école³. Tous les apprenants se situent entre les niveaux A1 et B1 du Cadre européen commun de référence.

L'analyse de la fréquence des formes dans l'input est basée sur trois sources données complémentaires, qui couvrent les trois principales variétés de français auxquelles les apprenants sont exposés.

a) Les données des interlocutrices natives du corpus pilote et de la conversation avec Lisa du corpus Lund. Il s'agit donc du français adressé à des apprenants (*foreigner talk*), dénué de constructions complexes tout en restant spontané et naturel.

b) Les données de quatre enseignants de FLE au lycée en Suède du corpus Flyman Mattsson (2003)⁴. Il s'agit de 39 enregistrements de leçons de français de durée variée pour un total d'environ 12 heures. Les enregistrements ont été faits dans trois classes à deux ans d'intervalle. Trois enseignants ont le suédois comme langue maternelle et une enseignante est locutrice native du français. Cette source de données permet d'avoir accès à l'input en classe (*teacher talk*), donc à l'input des apprenants du corpus pilote qui avaient appris le français dans leur jeunesse ainsi qu'à celui de l'apprenante Lisa du Corpus Lund.

c) Les données de *C-ORAL ROM* (2005), un DVD qui comprend 164 enregistrements (300'000 mots) de français parlé en France. On y trouve des monologues (par exemple un exposé), des conversations entre deux personnes ou plus dans des contextes privés et publics, des extraits d'émissions télévisées ainsi que des conversations au téléphone (y compris des répondants automatiques). Cette source de données permet ainsi d'avoir accès à l'input

³ Je remercie S. Schlyter d'avoir mis son corpus à ma disposition. Pour plus de détails sur ce corpus voir Granfeldt, 2003 ; Bartning & Schlyter, 2004.

⁴ Je remercie A. Flyman Mattsson d'avoir mis son corpus à ma disposition.

auquel les apprenants sont exposés lors de leurs séjours en pays francophone et plus particulièrement à celui de Karl du Corpus Lund.

5. Fréquence des formes de surface dans l'input

L'analyse des formes de surface dans l'input est basée sur 19 verbes de la conjugaison régulière en *-er*. Il s'agit de verbes couramment utilisés par les apprenants débutants. Le calcul de fréquence a été fait sur le nombre d'occurrences des verbes à l'infinitif, au participe passé, à l'imparfait (singulier et 3^e pluriel), et pour le présent de toutes les formes sauf celle de la 1^{ère} du pluriel, cette forme ayant une morphologie propre.

La compilation des formes chez les intervieweuses et les enseignants a été réalisée à l'aide du programme COMBO du logiciel CLAN (MacWhinney, 2000) qui présente les formes dans les énoncés et pour les données de C-ORAL-ROM du concordancier *contextes* fourni avec le DVD.

Le tableau 2 ci-dessous présente la fréquence en moyenne des formes courtes du présent et des formes longues se terminant en /e/ (*-er*, *-é*, *-ais*, *-ait*, *-aient* ; *-ez* au présent) dans l'input. Les résultats ont été triés à partir du pourcentage des formes courtes du présent.

	Σocc	% formes courtes	% formes longues en /e/
appeler (s')	61,2	72%	28%
habiter	21,2	72%	28%
trouver [état]	29,5	66%	34%
penser	61,3	62%	39%
aimer	45,3	60%	41%
rester	28,0	54%	46%
préférer	10,3	47%	53%
commencer	32,8	36%	64%
parler	92,0	36%	64%
jouer	25,6	33%	67%
chercher	17,6	33%	67%
travailler	41,2	26%	74%
regarder	35,7	25%	75%
écouter	16,3	24%	76%
manger	24,7	18%	82%
étudier	10,7	10%	90%
acheter	16,0	10%	90%
trouver [dyn]	33,0	9%	91%
voyager	9,6	7%	93%
visiter	5,8	4%	96%

Tableau 2 - Fréquence des formes dans l'input (moyennes)

On peut ainsi constater que les verbes pour lesquels la forme courte du présent est la plus fréquente en moyenne dans l'input analysé sont des verbes d'état et qu'à l'opposé, les verbes les moins fréquents à la forme courte du présent sont des verbes dynamiques, ce qui correspond à l'hypothèse au niveau de la distribution des formes dans l'input. La position du verbe *rester* en haut de la liste mériterait une analyse plus détaillée, notamment pour différencier les contextes statiques (*il reste du sucre*) des contextes dynamiques (*il reste à la maison*) dans l'utilisation de ce verbe. Pour le verbe *préférer* on peut constater que le résultat ne correspond pas à l'hypothèse, la forme en /e/ étant plus fréquente que la forme courte du présent. Le pourcentage de formes en /e/ pour ce verbe est essentiellement lié à la fréquence de la forme de la 2^e du pluriel. Sans celles-ci ce pourcentage descend en effet à 39%.

6. Données des apprenants

Les verbes analysés dans cette section sont les mêmes que pour l’input. Les calculs sont faits à partir du nombre d’occurrences de chaque verbe dans les contextes pour le présent, l’infinitif et le passé. Les répétitions ou imitations des formes exprimées par l’intervieweuse n’ont pas été comptées, du moment où elles n’ont pas été changées. Lorsque les apprenants répètent plusieurs fois la même forme verbale, seule la dernière a été prise en compte. De même lorsqu’il y a un changement entre la première forme et la deuxième forme du même verbe, seule la deuxième a été comptée, les apprenants comme les natifs devant avoir une chance de se corriger.

Les contextes pour l’infinitif correspondent au placement du verbe après une préposition (*pour parler*), après un verbe modal (*je veux parler*) ou semi modal (*j’aime parler*), l’auxiliaire du futur proche (*je vais parler*) et dans les autres constructions demandant l’infinitif, y compris dans une fonction de sujet (*jouer au golf n’est pas compliqué*).

Les contextes pour le présent correspondent à tous les contextes dans lesquels on attend une forme du présent en français, même si l’accord sujet-verbe n’est pas correct. La forme du présent a également été comptée comme correcte si l’apprenant passe au présent dans un récit du passé et s’y tient, ainsi que pour les événements du futur (*dans X temps je cherche du travail*) étant donné que le présent peut exprimer un futur relativement proche aussi en français natif.

Le tableau 3 ci-dessous présente la production des apprenants. Les verbes sont présentés dans le même ordre que dans le tableau 2 sur l’input. La première colonne après la liste des verbes contient la somme des formes produites à la forme courte (C) du présent (*parl-*) et à la forme longue (L) en /e/ (*parlE*). La colonne suivante présente le pourcentage de formes courtes du présent (correctement utilisées ou non), ce qui permet de voir la forme la plus utilisée pour chaque verbe par les apprenants. Les trois colonnes suivantes présentent d’une part la sur-utilisation de la forme courte du présent dans les contextes demandant une forme de l’infinitif et du passé et d’autre part la sur-utilisation de la forme longue dans un contexte demandant une forme du présent.

	Σ C + L	% C	Forme la + fréq apprenants	sur-ut C ds INF	sur-ut C ds PAS	sur-ut L ds PRE	Forme la + fréq input
appeler (s’)	32	97%	appelle	-	-	3%	appelle
habiter	79	85%	habite	60%	46%	5%	habite
trouver [état]	0	-	-	-	-	-	trouve
penser	31	97%	pense	100%	67%	0%	pense
aimer	49	100%	aime	-	-	-	aime
rester	9	56%	reste	50%	0%	20%	reste
préférer	8	100%	préfère	-	-	-	préfère
commencer	27	59%	commence	0%	0%	6%	commence
parler	119	70%	parle	17%	17%	8%	parle
jouer	16	13%	joue	0%	0%	75%	joue
chercher	9	44%	cherche	25%	0%	0%	cherche
travailler	72	69%	travaille	56%	56%	25%	travaille
regarder	26	38%	regarde	20%	13%	38%	regarde
écouter	13	23%	écoute	11%	0%	0%	écoute
manger	34	62%	mange	10%	57%	6%	mange
étudier	33	0%	étudie	-	-	100%	étudie
acheter	13	31%	achète	0%	14%	25%	achète
trouver [dyn]	10	30%	trouve	0%	33%	33%	trouve
voyager	22	23%	voyage	0%	18%	40%	voyage
visiter	23	0%	visite	-	-	100%	visite

Tableau 3 – Production des apprenants

Bien que l’on trouve des différences entre la forme la plus souvent produite par les apprenants et la forme la plus fréquente dans l’input, on peut voir qu’il y a une assez bonne

correspondance pour les premiers et les derniers verbes de la liste. On peut même noter que les apprenants ont uniquement produit le verbe *préférer* à la forme courte du présent, ce qui correspond à l'hypothèse d'un lien préférentiel entre les verbes d'état et la forme courte du présent. On peut également constater que la forme longue en /e/ n'a pratiquement pas été sur-utilisée pour les premiers verbes de la liste (de *s'appeler* à *préférer*), sauf pour le verbe *rester* dont on a déjà noté la double appartenance de verbe d'état et de verbe dynamique. Les apprenants ont uniquement utilisé ce verbe dans son sens dynamique (*il reste à la maison*).

Néanmoins, l'utilisation et la sur-utilisation de la forme courte du présent ne se limite pas à la production des verbes d'état. En effet, les verbes *commencer*, *parler*, *travailler* et *manger* ont avant tout été produits à la forme courte du présent par les apprenants, contrairement à ce qui était prédit pour les verbes dynamiques et contrairement à la fréquence dans l'input. On peut également constater que la plupart des verbes fréquents à la forme longue en /e/ chez les apprenants ont également été sur-utilisés à la forme courte du présent, en particulier dans les contextes demandant une forme du passé, à l'exception du verbe *travailler* pour lequel la forme courte du présent est clairement sur-utilisée dans les contextes demandant l'infinitif.

Même s'il existe un certain nombre de différences pour les verbes individuels, on peut constater la distribution de la forme courte du présent et de la forme longue en /e/ entre les verbes d'état et les verbes dynamiques, si l'on considère la moyenne des verbes pour chaque catégorie, comme le montre le tableau 4 ci-dessous.

	VERBES D'ÉTAT	VERBES DYNAMIQUES
% FORMES COURTES	96%	37%
% FORMES LONGUES	4%	63%

Tableau 4 - Distribution des formes courtes et longues chez les apprenants (moyennes)

La forme courte du présent s'avère ainsi plus fréquente pour les verbes d'état que pour les verbes dynamiques, alors que la forme longue en /e/ est plus fréquente avec les dynamiques qu'avec les verbes d'état, ce qui confirme l'hypothèse proposée.

7. Synthèse et discussion

Les recherches effectuées dans le cadre de la grammaire générative, en particulier par Prévost et White (Prévost & White, 2000 ; Prévost, 2004), ont montré que la variation des formes trouvées dans l'interlangue des apprenants n'était pas due à un déficit au niveau de la syntaxe, les apprenants ayant accès aux catégories fonctionnelles ainsi qu'aux traits abstraits de la finitude. La difficulté semble plutôt se situer au niveau de la production de la morphologie de surface. Ainsi les formes interlangagières utilisées par les apprenants débutants pourraient être caractérisées de formes par défaut, utilisées indépendamment de la fonction visée.

Une approche centrée sur l'usage de la langue a permis d'approfondir la question de l'utilisation de formes par défaut, proposant que les formes utilisées par les apprenants débutants soient influencées par la distribution des formes dans l'input.

L'étude présentée dans cet article a été faite sur 19 verbes réguliers en *-er*, d'abord analysés dans trois sources d'input, constitué de données correspondant aux principales variétés du français auquel les apprenants sont exposés, puis sur un corpus d'apprenants adultes, suédophones, débutants en FLE. A noter que la morphologie verbale du suédois comprend l'opposition entre la forme de l'infinitif et la forme courte du présent.

Les résultats ont permis de constater que la forme courte du présent (*habite*) était la forme la plus fréquente avec les verbes d'état dans l'input alors que la forme longue en /e/ (*parler*, *parlé*, *parlais/t/ent*, *parlez*) était plus fréquente avec les verbes dynamiques. L'analyse de la production des apprenants a permis de constater une distribution semblable à celle de l'input, même si l'on a pu observer quelques différences au niveau des verbes individuels.

Cette première analyse à partir d'un corpus de conversations a permis de vérifier l'hypothèse de la distribution des formes dans l'input et l'influence de celle-ci sur la production des formes par les apprenants. Dans une prochaine étape, le développement de tests expérimentaux permettra d'approfondir l'analyse, notamment sur la question de l'utilisation de formes par défaut par les apprenants. De tels tests permettront en effet de créer des contextes pour les formes courtes et les formes longues en /e/ pour tous les verbes, ainsi que de varier les contextes pour les verbes ayant une double appartenance sémantique, statique et dynamique, comme par exemple les verbes *trouver* et *rester*.

Cette étude n'est pas directement ancrée dans le domaine de la didactique du FLE, mais elle entend y contribuer à long terme, notamment au niveau de la possible influence de la langue utilisée en classe de FLE ainsi que de la didactique de l'enseignement de la morphologie verbale.

Références bibliographiques

- BARTNING, Inge & SCHLYTER, Suzanne (2004). Itinéraires acquisitionnels et stades de développement en français L2, *Journal of French Language Studies*, 14 : 281-299.
- BYBEE, Joan (1991). Natural morphology: the organization of paradigms and language acquisition, in *Crosscurrents in second language acquisition and linguistic theories*, Huebner, Thom & Ferguson, Charles (Eds). Amsterdam/Philadelphia : John Benjamins, 67-91.
- BYBEE, Joan (1985). *Morphology. A study of the relation between meaning and form*. Amsterdam/Philadelphia : John Benjamins.
- C-ORAL-ROM (2005). *Integrated Reference Corpora for Spoken Romance Languages*. Cresti, Emanuela & Moneglia, Massimo (Eds). Amsterdam/Philadelphia : John Benjamins.
- ELLIS, Nick C. (2002). Frequency Effects in Language Processing. A Review with Implications for Theories of Implicit and Explicit Language Acquisition, *Studies in Second Language Acquisition*, 24: 143-188.
- FLYMAN MATTSSON, Anna (2003). *Teaching, Learning, and Student Output. A study of French in the classroom*. Thèse de doctorat, Université de Lund.
- GRANFELDT, Jonas (2003). *L'acquisition des catégories fonctionnelle. Etude comparative du développement du DP français chez des enfants et des apprenants adultes*. Thèse de doctorat, Université de Lund.
- HAWKINS, Roger (2004). The contribution of the theory of Universal Grammar to our understanding of the acquisition of French as second language, *French Language Studies*, 14: 233-255.
- KIHLSTEDT, Maria (1998). *La référence au passé dans le dialogue. Etude de l'acquisition de la temporalité chez des apprenants dits avancés de français*. Thèse de doctorat, Université de Stockholm.
- MACWHINNEY, Brian (2000). *The CHILDES project : Tools for analyzing talk*. Mahwah, NJ : Lawrence Erlbaum Associates.
- MITCHELL, Rosamond & MYLES, Florence (1998). *Second Language Learning Theories*. London : Arnold.
- PRÉVOST, Philippe (2004). Morphological variation in early adult second language French. A cross-sectional study, in *EUROSLA Yearbook 4*, Foster-Cohen Susan et al. (Eds). Amsterdam/Philadelphia : John Benjamins, 147-175.
- PRÉVOST, Philippe & WHITE, Lydia (2000). Missing Surface Inflection or Impairment in second language acquisition? Evidence from tense and agreement, *Second Language, Research* 16/2 : 103-133.
- THOMAS, Anita (2006). Les formes du type *je parle* dans l'interlangue d'apprenants du FLE et la fréquence dans l'input, *PERLES*, 23, Université de Lund.
- VENDLER, Zeno (1957). Verbs and Times, *Psychological Review*, 66: 143-60.

SOCIOLINGUISTIQUE ET PLURILINGUISME

LA STANDARDISATION DE L'AMAZIGHE (BERBÈRE) AU MAROC : ENJEUX ET RÉCEPTION AUPRES D'ENSEIGNANTS

Myriam ABOUZAÏD

Myriam.Abouzaid@u-grenoble3.fr

Laboratoire Lidilem

Université Stendhal Grenoble III

BP 25

38040 Grenoble cedex 9

Abstract : In Morocco, the process of officially recognising the Amazigh language (Berber) started in 2001. This formerly minority-status language is now being standardized. This paper describes the main choices made in corpus planning of Amazigh and analyses how the newly standardized language is perceived by teachers of Amazigh.

1. Contexte de l'étude

La langue amazighe¹ est considérée comme la plus ancienne langue d'Afrique du Nord. Elle est présente dans tout le Maghreb, mais c'est au Maroc que se trouve la population amazighophone la plus importante en nombre, les estimations allant de 30% à 50% de la population marocaine. Cette langue s'inscrit dans un paysage linguistique où elle cohabite avec l'arabe dialectal marocain, l'arabe standard et le français. Le Maroc, à l'instar de tous les pays arabes, présente une situation de diglossie². À y regarder de plus près, on s'aperçoit que cette diglossie est « enchâssée », selon les termes de Calvet (1987 : 47). C'est-à-dire que la langue amazighe est triplement minorée, se situant tout en bas de l'échelle du marché linguistique marocain, après l'arabe standard, le français, et même après l'arabe dialectal.

L'année 2001 marque un véritable tournant dans la politique linguistique marocaine, avec la création, par le Roi Mohamed VI, de l'Institut Royal de la Culture Amazighe (désormais « IRCAM »). Jusqu'alors, l'amazighe ne bénéficiait d'aucune reconnaissance officielle. Cet institut a pour mission d'aménager la langue afin qu'elle puisse être enseignée et jouer un rôle véritable à l'échelle nationale. La standardisation, première étape de l'aménagement linguistique, représente un travail de grande envergure puisque l'amazighe, langue dite « de tradition orale », a évolué, au cours des siècles, en une immense variété de parlers. Au Maroc, il est possible d'identifier trois aires dialectales majeures : au Nord-Est, le *tarifite* (ou *rifain*), au Centre, le *tamazighte*, et au Sud, le *tachelhite*. Précisons que si la grande majorité des locuteurs et des linguistes s'accordent à dire que ces trois variétés relèvent de la même langue, il n'en demeure pas moins que, à l'heure actuelle, l'intercompréhension se révèle parfois difficile entre amazighophones géographiquement éloignés.

À partir de cette brève description de la situation linguistique de l'amazighe au Maroc, il semble légitime de s'interroger sur la réception de la langue nouvellement unifiée. Plus précisément, il est possible de formuler cette interrogation comme suit : l'unification des variétés dialectales amazighes peut-elle aboutir à la création d'une diglossie amazighe, avec

¹ Au Maroc, le terme générique désignant la langue berbère est désormais « amazighe ». Cependant, on trouve aussi « tamazight(e) » employé dans un sens générique.

² Situation relativement stable dans laquelle deux variétés d'une même langue existent côte à côte dans une communauté, l'une d'entre elles étant reconnue comme « haute » et l'autre « basse » (Ferguson, 1959)

une langue « haute », la langue standard, unifiée, qui serait réservée à l'écrit et aux domaines formels, et plusieurs variétés – les différents vernaculaires – qui demeureraient « basses » ?

Afin d'apporter des éléments de réponse à cette question, nous avons effectué une enquête de terrain auprès de quinze instituteurs qui, depuis 2003, doivent enseigner l'amazighe à leurs élèves du cycle primaire. L'approche de cette recherche est donc essentiellement sociolinguistique. Des entretiens semi-directifs ont été menés à la fois dans des zones géographiques fortement amazighophones et dans des zones majoritairement arabophones. Nous avons interrogé ces enseignants sur leur expérience personnelle d'enseignement de l'amazighe, et en particulier sur leur perception de la langue du manuel pédagogique. Nous avons également rencontré six linguistes de l'IRCAM, ce qui nous a permis de mieux cerner le processus de codification de la langue et ses enjeux. Les informations ainsi recueillies font l'objet de la partie 2 de cet article. Quant aux entretiens avec les enseignants, ils ont été soumis à une analyse de contenus (thématique), laquelle a mis au jour des représentations sociales à l'égard de la langue nouvellement standardisée. Ces représentations sont présentées dans la partie 3.

2. Les décisions officielles en matière d'aménagement du corpus et leurs enjeux

2.1. Une gestion démocratique de la langue standard

Dans de nombreux cas, l'aménagement d'une langue, passe par l'identification d'une variété (souvent socialement dominante) qui sert de base à la standardisation. C'est la manière la plus « simple » de standardiser, car elle présente un degré d'intervention moindre. Pour l'amazighe, il est impossible d'identifier la variété remplissant de la meilleure façon le rôle de référent. En effet, de l'absence de « tradition écrite » résulte l'absence d'une langue-mère attestée dont seraient issues les variantes linguistiques actuelles. De surcroît, la démarche consistant à sélectionner une variété de base est écartée d'emblée par les linguistes de l'IRCAM, et ce, pour des raisons idéologiques. Il faut que la langue standard épouse la réalité sociolinguistique du pays. On constate que l'aménagement de l'amazighe se situe dans l'optique de la *koinè* grecque, soit une langue créée sur la composition de différents dialectes. Il s'agit d'un processus comparable à l'élaboration de la « langue polynomique » corse qui « rend compte de la diversité dialectale, lui confère un caractère légitime et ne hiérarchise pas les différentes normes en présence » (Comiti & Di Meglio, 1999 : 63). Les aménageurs de l'amazighe sont donc confrontés à une tâche extrêmement complexe avec comme règle d'or de préserver « l'unité dans la diversité ».

2.2. Une standardisation progressive et convergente

Les spécificités de l'amazighe précédemment décrites, ainsi que la politique des aménageurs impliquent une standardisation progressive, convergente, en deux étapes. La première s'effectue sur le plan intra-géolectal, c'est-à-dire à l'intérieur des trois grandes aires dialectales. La deuxième se situe au niveau inter-géolectal, c'est-à-dire à l'échelle du Maroc (Ameur *et al.*, 2004 : 18). S'il apparaît logique de considérer que l'aménagement du corpus est un préalable à l'aménagement du statut d'une langue, il faut toutefois noter que la standardisation se fait également, et en grande partie, *par* l'enseignement. Les deux étapes de l'aménagement linguistique se retrouvent donc dans le cadre scolaire. C'est pourquoi il existe une régionalisation du manuel pédagogique *Tifawin* (IRCAM, 2003 & 2004). Le manuel de 1^{re} année est ainsi décliné en trois versions identifiables par des couleurs différentes. Le contenu est identique, mais la langue demeure fidèle, autant que possible, aux spécificités respectives du *tarifite*, du *tamazighte* et du *tachelhite*. Le manuel de 2^e année se présente sous une version unique, mais dans lequel le système des couleurs est maintenu.

Certaines pages reflètent chaque variété séparément, et quelques parties sont communes aux trois. Ainsi, l'apprenant est exposé petit à petit aux deux autres variétés. L'objectif final est de parvenir, en quelques années, à un manuel unique présentant une langue véritablement unifiée.

2.3. Les principaux domaines d'action des aménageurs

2.3.1. Le choix d'un alphabet officiel

Dans le cas d'une langue « de tradition orale », la sélection d'un système d'écriture apparaît comme la toute première étape de la standardisation. Au Maroc, officialiser un alphabet précis a permis de mettre fin à l'« anarchie graphique » qui caractérisait la situation de l'amazighe. En effet, jusqu'alors, les systèmes de notation employés pour transcrire la langue relevaient, à divers degrés, de trois graphies : la graphie latine, la graphie araméenne (arabe) et la graphie *tifinaghe*. Cette dernière constitue le système d'écriture spécifique à l'amazighe, dont l'ancêtre est désigné par les linguistes sous le nom d'alphabet *libyco-berbère*. Cependant, son usage a toujours été très restreint.



Figure 1 - Aperçu de l'alphabet *tifinaghe*³

En 2003, une décision royale a opté pour l'alphabet *tifinaghe* comme graphie officielle de la langue amazighe au Maroc, sur les recommandations de l'IRCAM. En amont comme en aval de cette décision, les réactions ont été vives. On a même parlé de « guerre des graphies ». Les arguments des uns et des autres relèvent à la fois de la pertinence scientifique attribuée à tel ou tel système, et des préférences idéologiques sous-jacentes. Pour résumer très schématiquement, les partisans des caractères arabes souhaitent voir s'inscrire l'amazighe dans le contexte arabo-musulman national. Les tenants de la graphie latine, eux, prônent une ouverture sur la culture universelle et moderne. Ils rappellent également qu'elle est le support utilisé majoritairement dans la documentation scientifique existante. La graphie *tifinaghe* incarne à la fois une certaine neutralité et une dimension historique et authentique. Le choix officiel, on le devine, fut très controversé et politisé.

Nous avons pu constater, sur le terrain et y compris chez les linguistes de l'IRCAM, que l'alphabet latin n'est pas véritablement abandonné au profit du *tifinaghe*. Si ce dernier est retenu pour l'enseignement scolaire, il est des domaines où la graphie latine perdurera, de façon transitoire ou non. Il se pourrait donc qu'une double graphie soit en train de s'instaurer au Maroc (*tifinaghe* et alphabet latin).

2.3.2. La gestion de la variation phonétique

Les choix effectués dans le passage à l'écrit découlent des choix stratégiques généraux d'unification de la langue. Il a donc fallu aboutir à un système graphique qui transcende les manifestations orales. La standardisation de la morphosyntaxe passe par une écriture dite « grammaticale » (Taïfi, 2004 : 32), ou « étymologique », ou encore « à base de dérivation ». Ainsi, le plus souvent, pour retenir une forme standard, il faut identifier les racines qui

³ Illustration empruntée à Atanane Oulahyane (http://www.asays.com/article.php3?id_article=346).

peuvent être dissimulées par des changements phonétiques ou des assimilations. Autrement dit, les aménageurs opèrent une reconstruction diachronique, ce qui permet de dériver les formes dialectales existantes, mais aussi, et surtout, d'en supposer d'autres possibles (Taïfi, 2004). Il s'agit ici d'une démarche qualifiée de « proto-berbère ». La difficulté principale réside dans le fait que, en l'absence d'identification d'une langue mère, il est particulièrement délicat de décider de ce qu'est véritablement le « proto ». Un des effets de cette écriture consiste en une distance parfois importante entre la graphie et la prononciation d'un terme. Il existe cependant, dans cette standardisation, une autre démarche qui, *a priori*, peut être concurrente de celle décrite précédemment : c'est la démarche « pan-berbère ». En effet, le critère de l'extension géographique d'une forme de prononciation est également pris en compte. Autrement dit, il a fallu statuer sur ce qui relève du local et ce qui est généralisé à la plupart des variétés régionales. Lorsque la forme est généralisée, elle a de fortes chances d'être retenue comme standard. Le travail des linguistes réside donc dans l'évaluation de différents critères de sélection, avec, comme souci permanent, celui de la pertinence, c'est-à-dire du caractère discriminant des orthographes retenues.

2.3.3. La gestion de la variation lexicale

Les linguistes s'accordent à dire qu'il existe environ 50% de lexique commun aux différentes variétés dialectales. Cela signifie que l'amazighe connaît une diversité extrême sur le plan lexical. L'on devine que cette richesse terminologique peut poser de sérieuses difficultés lorsqu'est adoptée une gestion démocratique de l'aménagement. Dans la mesure du possible, ces variantes lexicales sont traitées comme des synonymes. Cette diversité est donc exploitée dans le but d'affiner les significations, d'enrichir la stylistique de la langue, tout en comblant ses manques lexicaux. Il s'agit du processus de redistribution sémantique. Cette démarche se révèle particulièrement utile dans le domaine de l'abstrait, où, de façon générale, il manque des termes.

2.3.4. La gestion des emprunts

L'amazighe présente une forte influence des langues avec lesquelles il a été en contact, notamment le français, l'espagnol, et surtout l'arabe. Ainsi, le *tarifite* présente de nombreux emprunts à l'espagnol et à l'arabe, tandis que les dialectes du Sud sont généralement considérés comme plus conservateurs. À titre d'exemple, seul le *tachelhite* a conservé les chiffres amazighes ; la plupart des parlers des autres variétés ont depuis longtemps adopté les chiffres arabes. Dans ce cas précis, l'amazighe standard réintègre les chiffres amazighes, qui, anciennement, étaient utilisés par tous. Pour « retrouver » un terme amazighe dont l'usage s'est fait rare, les recherches des linguistes étudient les parlers proches, dans un premier temps, puis les dialectes éloignés, et, en dernier recours, le kabyle (variété amazighe principale d'Algérie) et le touareg (variété amazighe présente en Algérie, en Libye, au Niger, au Mali et au Burkina Faso). On comprend ainsi que l'élaboration de néologismes constitue le dernier recours des linguistes, en cas de vide dans le répertoire amazighe pour désigner un référent.

De façon générale, dans la gestion de l'aménagement linguistique, les mots d'ordre semblent être « modération » et « prudence ». Les linguistes sont bien conscients du risque d'épuration de langue, et de l'écueil qui ferait aboutir leurs travaux à une « langue de laboratoire » méconnaissable de ses locuteurs. Malgré cette prudence, à partir des éléments que nous venons de décrire, il semble légitime de se demander si une situation de diglossie amazighe n'est pas effectivement en gestation ici. Il semble en effet que plusieurs conditions soient réunies. Afin d'éclairer davantage cette hypothèse, il apparaît nécessaire d'observer les discours des enseignants amazighophones que nous avons rencontrés.

3. La réception du travail de standardisation auprès d'enseignants d'amazighe

Précisons qu'il s'agit ici d'une étude qualitative et non quantitative. À partir de quinze entretiens semi-directifs, nous avons retenu des paroles⁴ estimées représentatives des discours des instituteurs rencontrés.

3.1. L'alphabet tifinaghe

Au vu des débats idéologiques ayant entouré la sélection de la graphie officielle, on aurait pu s'attendre à retrouver, dans le discours des enseignants, une illustration des différents arguments en jeu. Or, nous n'avons pu recueillir que très peu de réactions face au choix de la graphie officielle. Trois grands types de réaction sont à relever parmi les instituteurs. Le premier concerne la mise en avant de l'aspect pratique de l'alphabet latin, face au *tifinaghe* :

« Je suis d'accord que c'est un héritage culturel, mais est-ce qu'il va permettre un épanouissement de cette langue? Je crois que non. Le latin aurait permis que les autres nous lisent »

Ici, le *tifinaghe* semble constituer un obstacle à la propagation de la langue. Cette opinion est d'ailleurs très souvent partagée par les enseignants militant pour la reconnaissance de la langue-culture amazighe dans un cadre associatif.

Ensuite, les principaux arguments avancés en faveur de l'alphabet *tifinaghe* sont sa simplicité, son authenticité et son historicité :

« C'est facile à dessiner pour les enfants, et puis c'est un art marocain. Tout le monde connaît le tfinaghe dans l'art, dans les grottes... »

On notera l'emploi du terme « dessiner » qui suggère l'aspect ludique, car inédit, conféré à cette graphie. Enfin, nombreuses ont été les réactions « neutres », comme ici :

« L'amazighe, c'est une langue comme toutes les autres. Les lettres sont nouvelles? Il faut les accepter. Toutes les critiques... non (...) Et puis l'alphabet amazighe est très facile. Il n'y a pas de problème »

En somme, les consignes étant ce qu'elles sont, l'adoption du nouvel alphabet n'est guère remise en question par le corps enseignant.

3.2. Réactions ambivalentes face au lexique standard

Lorsque nous avons questionné les instituteurs au sujet du vocabulaire des manuels pédagogiques, ils ont, dans l'ensemble, montré une attitude mitigée face au traitement de la variation. Les deux extraits d'entretiens suivants font état d'un lexique standard qui se révèle déroutant pour certains enseignants (ici, des enseignants tamazightophones) :

« Le problème, c'est la langue standard (...) On est amazighophones mais on a des difficultés à comprendre certains mots du livre »

« On dirait que la langue est celle du Souss⁵, au niveau du vocabulaire surtout, et c'est un problème car on ne connaît pas tous les mots. Il nous faudrait un dictionnaire pour comprendre tout »

⁴ L'intégralité des entretiens dont sont extraites ces paroles apparaît dans notre mémoire : ABOUZAIID, Myriam (2005) « L'aménagement de l'amazighe au Maroc : enjeux et réception auprès des enseignants ». Mémoire de Master 2 Recherche en Sciences du langage, sous la direction de Billiez Jacqueline, Laboratoire Lidilem, Université Stendhal, Grenoble 3.

⁵ Souss : région du Sud du Maroc, où est parlée la variété *tachelhite*.

Une certaine distance est ainsi établie entre la langue du manuel et la langue première des enquêtés. Sans aller jusqu'à rejeter les termes standard, ces enseignants évoquent néanmoins cette distance en termes de « problème ». Précisons que ces paroles illustrent essentiellement la vision des enseignants des zones fortement amazighophones où la langue standardisée vient se « confronter » au(x) vernaculaire(s) amazighe(s) des élèves. Dans les régions majoritairement arabophones, les réactions négatives face à la langue standard ont été bien moins fréquentes.

Il semble donc qu'une partie des enseignants rencontrés soit quelque peu déstabilisée par le lexique standard des manuels pédagogiques. Cependant, il est à noter que lorsque le travail de standardisation porte sur le remplacement des emprunts (principalement à l'arabe), un accueil relativement bon est réservé aux néologismes⁶. Ainsi, les nouveaux termes semblent mieux acceptés que les termes « rapatriés » d'autres régions, comme le suggèrent les deux extraits suivants :

« Par exemple, pour dire « livre », on dit ktab. C'est un mot arabe. On utilise beaucoup de mots arabes. L'IRCAM a remplacé tous ces mots par des mots amazighes. Pour ktab on doit dire adliss »

« Maintenant je suis contente quand tous mes mots sont amazighes. Et je suis heureuse parce que je ne parle plus moitié-moitié... tous mes mots sont amazighes. Même s'ils sont nouveaux, je suis contente »

On remarque ici que le travail sur la langue vient atténuer ce que Calvet (1999 : 168) nomme l'« insécurité formelle », c'est-à-dire le sentiment de ne pas parler comme il faudrait parler. Il apparaît donc que lorsque les mots nouveaux s'inscrivent dans une optique de chasse aux emprunts superflus, le travail d'adaptation à la langue standard s'en trouve facilité.

3.3. Accueil positif des normes linguistiques

Outre cette satisfaction de voir la langue « gagner en amazighité », il est apparu que l'officialisation des normes linguistiques agit clairement en faveur de la langue standard. Les deux extraits suivants montrent que l'apparition de ces normes rehausse fortement le prestige accordé à la langue :

« Avant, je n'avais jamais pensé que la langue amazighe avait des règles elle aussi. Mais maintenant, j'ai compris que oui. Et ce sont les mêmes règles partout »

« Il y a des règles d'orthographe en amazighe, la segmentation est très importante. Ecrire un mot, c'est très facile, plus facile qu'en français, mais une phrase, là, c'est plus compliqué (...) C'est une langue riche, qui a des règles, il ne suffit pas de... »

Cette enseignante souligne, à plusieurs reprises, l'existence de règles grammaticales et orthographiques en amazighe, comme si avant cette prise de conscience, elle estimait que sa langue était une « langue sans règles », une langue « anarchique ». Il semble donc que les règles d'usage, les régularités repérables dans les discours aient été considérées comme inexistantes jusqu'à ce qu'elles soient reconnues officiellement comme des normes. Il s'agit ici d'une conséquence de la minoration que l'amazighe a toujours subie. Mais ici, plus encore que son officialisation, c'est la mise au jour de cette grammaire intériorisée par ses

⁶ Nous ne faisons pas, ici, la distinction entre néologismes « purs » (mots nouvellement forgés par les aménageurs de la langue) et termes « réhabilités » (qui peuvent être inconnus dans certaines régions, mais d'usage courant dans d'autres), l'important étant la nouveauté qu'ils présentent pour les enquêtés.

locuteurs qui paraît agir en faveur de la langue et lui redonner de la crédibilité en tant que langue.

Ainsi, la réception des différents éléments de la standardisation s'avère quelque peu ambiguë : elle traduit à la fois une satisfaction de voir la langue s'enrichir et se doter de normes, mais elle laisse entrevoir un certain scepticisme initial face à l'ampleur de la nouveauté lexicale. En outre, il est arrivé fréquemment de retrouver cette ambivalence chez une même personne.

3.4. *Un terrain propice à une nouvelle diglossie ?*

Naturellement, nous avons cherché à savoir quelle place occupe la langue nouvellement standardisée dans ce que l'on pourrait appeler le « marché linguistique intérieur » des locuteurs. Autrement dit, prend-elle, à leurs yeux, la place des vernaculaires ? Ou bien vient-elle se superposer à ces parlers telle la variété « haute » dans un schéma de diglossie, venant occuper une place jusqu'alors vide ?

Sur ce point, plusieurs enquêtés ont spontanément dressé un parallèle avec la langue arabe :

« Pour la langue standard, c'est l'équivalent du classique »

« C'est comme pour l'arabe... Est-ce qu'un peuple parle la langue classique ?

Non ! Donc ce n'est pas un problème »

On constate ici l'impact de l'environnement linguistique préexistant, c'est-à-dire la diglossie arabe, laquelle diglossie semble véritablement incarner un modèle de référence. Dans les extraits suivants, la « langue de l'école » réfère à la langue « haute » :

« On montre aux élèves que c'est la langue amazighe moderne, que c'est la langue de l'école »

« La langue que nous devons enseigner n'est pas la langue de la rue, c'est une langue qui doit servir à raisonner. La langue de l'école doit être une langue aménagée, qui va servir, plus tard, à des buts plus élevés »

La dichotomie « langue de la rue » / « langue de l'école » est ainsi établie comme un postulat qui neutralise en amont tout discours visant à remettre en cause la langue standard. Il est ici question de registres différents dont on ne saurait confondre l'usage. Il semblerait, que selon les instituteurs, « la langue de la rue » ne permette pas de « raisonner ». En outre, souligner ainsi le manque d'équipement de la variété « de la rue » conduit non seulement à distinguer les domaines d'usage de cette langue, mais également à les hiérarchiser. Autrement dit, la rue en bas, l'école en haut. Il semble donc que la légitimation de la forme standard ne puisse faire l'économie du rabaissement de sa congénère.

Peu de gens, au Maroc, remettent en question la légitimité du couple arabe classique/arabe dialectal. La langue standard viendrait donc cohabiter avec les variétés parlées tout aussi légitimement que dans le cas de l'arabe. Notons également que la comparaison avec la langue officielle du pays, si elle est logique et naturelle, n'en est pas moins révélatrice d'une situation conflictuelle. Pour certains enquêtés, la langue arabe semble bien être un concurrent de l'amazighe, voire une menace. La langue amazighe standard apparaît ainsi comme un remède à cette concurrence déloyale. Si un profil diglossique de la langue amazighe se construit à l'image de celui de la langue arabe, cela n'apparaît pas comme un inconvénient mais bien comme un avantage.

4. Conclusion

D'après l'ensemble des témoignages recueillis, il apparaît que la nécessité de parvenir à une langue standardisée et unifiée n'est pas remise en question. La standardisation est perçue comme un investissement à long terme, et les enseignants – qui sont à la fois locuteurs et

acteurs de cette standardisation – semblent conscients des efforts d’adaptation à fournir pour que cette langue standard s’implante véritablement dans le paysage linguistique marocain. Nous avons montré que des contradictions subsistent toutefois à l’égard de la langue unifiée: elle incarne à la fois une aspiration et un objet de méfiance. Aspiration, car l’établissement de normes engendre un fort rehaussement du prestige de la langue. Ceci constitue un atout majeur à l’égard de l’acceptation de la langue standard. Méfiance, car lorsque le parler quotidien diffère grandement de la langue unifiée, la distance est parfois perçue comme brutale.

Enfin, le trait saillant de cette étude réside dans la constatation suivante : si, comme nous le postulons, la langue unifiée semble susceptible d’acquérir de nombreux traits d’une variété « haute » dans un schéma de diglossie, les réactions de nombreux enseignants révèlent que la diglossie arabe préexistante a déjà préparé le terrain à cette éventualité et que, loin de constituer un handicap, une telle évolution ne ferait que valoriser le nouvel idiome.

Références bibliographiques

- AMEUR, Meftaha & BOUMALK, Abdallah (Dir) (2004). *Standardisation de l’amazighe*, Actes du séminaire organisé par le Centre de l’Aménagement Linguistique à Rabat, 8-9 décembre 2003, Publication de l’Institut Royal de la Culture Amazighe, Série : Colloques et séminaires.
- CALVET, Louis-Jean (1987). *La guerre des langues et les politiques linguistiques*, Paris : Payot.
- CALVET, Louis-Jean (1999). *Pour une écologie des langues du monde*, Paris : Plon.
- COMITI, Jean-Marie & DI MEGLIO, Alain. (1999). « La langue corse dans l’enseignement : norme(s) et fonctionnements » in *Les langues régionales. Enjeux sociolinguistiques et didactiques*, Dabène, Louise (coord.). Grenoble : Presses Universitaires de Grenoble, 61-74.
- FERGUSON, Charles (1959). Diglossia. *Word*, 15: 325-340.
- IRCAM (2003). *Tifawin a tamazight 1, adlis n unlmad* (manuel pédagogique de l’élève), Casablanca : Afrique Orient.
- IRCAM (2004). *Tifawin a tamazight 2, adlis n unlmad* (manuel pédagogique de l’élève), Rabat : Publication OKAD.
- TAÏFI, Miloud (2004). Si les Berbères ne s’entendent pas, qu’ils s’écrivent ! Pour une écriture grammaticale du berbère à usage didactique in *Standardisation de l’amazighe*, Ameer, Meftaha & Boumalk, Abdallah (Dir), Rabat : IRCAM, 30-43.

REPRÉSENTATIONS DE MIGRANTS ORIGINAIRES DU MAGHREB : DISTINCTION DES VARIÉTÉS DIGLOSSIQUES PAR LES REPRÉSENTATIONS SUR LA PRODUCTION ET LA RÉCEPTION

Luc BIICHLÉ

luc.biichle@libertysurf.fr

Université Stendhal Grenoble III

UFR des Sciences du Langage

BP 25 - 38040 Grenoble cedex 9

Abstract : According to scientists, the main characteristic of diglossia can be found in the functional separation that exists between high and low varieties. The purpose of the following article is to confirm this separation but in terms of social representations of migrants from the Maghreb in France.

Dans le cadre d'une thèse consacrée aux langues et parcours d'intégration d'immigrés maghrébins de France, nous nous sommes intéressé au répertoire verbal de ces derniers et, notamment, au rapport qui existe entre les variétés dites « haute » et « basse » dans le concept de diglossie.

1. Bref aperçu de la situation sociolinguistique au Maghreb

Le Maghreb est constitué de trois pays (Tunisie, Maroc, Algérie) dont la langue officielle est l'arabe. Sur cet ensemble géographique, coexistent plusieurs familles de « langues », « dialectes » ou « variétés » (pour ne pas prendre parti puisque « nommer une langue, c'est déjà prendre position » (Tabouret-Keller, 1982 : 27)) que l'on peut regrouper de manière lapidaire : l'arabe qui est la langue de la majorité des locuteurs, le berbère qui concerne un nombre important de locuteurs et le français. Le système linguistique du Maghreb est ce que l'on appelle, selon la définition de Ferguson (1959), une *diglossie* que l'on peut définir rapidement comme la coexistence :

- d'une variété haute de langue arabe qualifiée de littéraire, standard ou classique, apprise à l'école, lue, écrite, parlée exclusivement en contextes formels (H), qui se trouve également être la langue du Coran, ce qui lui confère un statut très particulier (Romaine, 1989 : 37 ; Calvet, 1999 : 240), celui de langue prestigieuse, noble et sacrée,
- avec des variétés qualifiées de basses, dialectes, variétés et vernaculaires, arabes et berbères (B) exclusivement utilisées en contextes informels, dans la vie courante.

Il est possible de reprocher à cette présentation assez « manichéenne » de ne pas prendre en compte les espaces intermédiaires ou le continuum entre les variétés puisque certains auteurs (Laroussi, 1993 ; Taleb Ibrahim, 1995) attestent de la présence d'un arabe médian ou moyen entre H et B ainsi que de différences entre arabe littéraire et classique, tout comme d'autres auteurs attestent de la présence d'autres variétés de français (Caubet, 1998 ; Abbassi, 1977). Mais il faut ajouter à cette première répartition diglossique le français, langue de l'ancien colonisateur, qui est toujours présent dans beaucoup de domaines comme l'administration ou l'enseignement supérieur par exemple, et qui peut être considéré comme une seconde variété haute. C'est d'ailleurs cette présence du français qui fait dire à certains auteurs (Romaine, 1989 : 34 ; Calvet, 1999 : 231) qu'il s'agit d'une triglossie plutôt que d'une diglossie. Cependant, que l'on parle de di- ou de triglossie, le critère majoritairement retenu par les chercheurs est la séparation fonctionnelle des variétés : H en contexte formel et B dans la vie

courante. Par ailleurs, si la variété haute est généralement celle de l'écrit, certaines variétés d'arabe dialectal sont en voie de standardisation et sont maintenant aussi utilisées à l'écrit pour les lettres et nouvelles. Il en va de même pour le(s) berbère(s) qui est (sont) également en voie de standardisation et qui dispose(nt) aussi d'écritures (Ameur & Boumalk, 2004 ; Chaker, 2002). Enfin, concernant les variétés basses, on peut selon certains auteurs et locuteurs, opérer des distinctions entre arabe marocain, algérien et tunisien (Caubet, 1998 ; Laroussi, 1993 ; Benjelloun, 2002 ; Embarki, 2004 ; Allal, 2004), différences qui sont, en outre, attestées par l'existence de cours d'arabe marocain dispensés par l'INALCO (Institut National des Langues et Civilisations Orientales), par l'épreuve facultative d'arabe dialectal (5 variétés dialectales au total) proposée au baccalauréat français, ou encore par les différentes variétés d'arabe proposées dans les ELCO (Enseignement des Langues et Cultures d'Origine) des écoles primaires françaises. Nous nous appuyerons donc, dans le cadre de cet article, sur les notions de di- et/ou de triglossie (présence du français en plus en variété haute) sans tenir compte des réserves précédemment émises quant à la réalité de la séparation fonctionnelle. En revanche, nous prendrons en compte les distinctions entre les arabes marocain, tunisien *et* algérien ainsi qu'entre les différentes « langues », « dialectes » ou « variétés » berbères.

2. Le recueil de données

Les données sur lesquelles se fonde cette étude sont le résultat d'entretiens semi-directifs enregistrés (individuels et très majoritairement en français), effectués auprès d'un groupe de 105 personnes, toutes originaires du Maghreb, pour la plupart fraîchement arrivées en France (70% de primo-arrivants, moins de trois ans en France [critère élargi] et 30% d'ex-primo-arrivants, plus de trois ans en France), et rencontrées dans 13 organismes de formation de la région Rhône-Alpes au sein desquels se sont déroulés les entretiens (en « tête-à-tête » dans un espace isolé). Les trois nationalités maghrébines (algérienne, 60% ; tunisienne, 15% ; marocaine, 24%) sont représentées tout comme le sont la plupart des locuteurs des différents parlers berbères (kabylo, tachelhit, tamazight, tarifit, chaoui), locuteurs qui demeurent cependant minoritaires au sein du groupe enquêté. Les femmes sont nettement majoritaires (63%) et la fourchette d'âge s'étend de 16 à 60 ans avec une majorité d'enquêtés âgés de 16 à 30 ans (49%). La plupart des personnes relèvent de l'immigration dite « familiale », 82% d'entre elles sont sans emploi (majoritairement les femmes), 55% ont été scolarisées, 16% un peu et 29% pas du tout (très majoritairement les femmes). Il ne s'agit donc pas d'un échantillon représentatif mais plutôt de la « photo » d'une tranche de population de la région Rhône-Alpes, rencontrée sur des lieux de formation en 2004. Au cours de ces entretiens, basés sur la méthodologie d'une enquête précédente (Biichlé, 2003), les personnes ont été interrogées, entre autres (parcours migratoire, situation sociale, réseau social, difficultés rencontrées, aides éventuelles, etc.), sur les langues, dialectes ou variétés qu'elles parlaient ou comprenaient et ce, à l'aide de questions ouvertes et fermées. Ces questions d'ordre « linguistique » visaient à mettre en évidence les représentations des personnes sur leur répertoire verbal et, plus particulièrement, sur leur rapport à la langue du pays d'immigration, « le » français. Une première distinction, clairement partagée par les enquêtés et confirmée par nos travaux précédents, a été opérée entre production et réception (« parler » et « comprendre »), et une seconde, émanant plus directement du chercheur, a pu être établie entre les questions ouvertes et les questions fermées ; les résultats obtenus variant très souvent d'un type de questionnement à l'autre. Dans une précédente étude (Biichlé, 2003), ces écarts ont atteint parfois 40%... Nous avons, par conséquent, opté pour deux types de questionnement espacés, questions fermées en début d'entretien *vs* questions ouvertes à la fin, au sein desquelles nous avons dissocié la production de la réception :

- Questions fermées (début d'entretien) : Parlez-vous « telle » langue ? Comprenez-vous « telle » langue ?
- Questions ouvertes (fin d'entretien) : Quelles sont les langues que vous parlez ? Quelles sont les langues que vous comprenez ?
- Les diverses réponses ont été réparties en trois types : les « oui », les « non » et les « un peu », dernière catégorie moins précise qui regroupe les « un petit peu », les « pas beaucoup », les « peu », etc.

2.1. Les questions fermées

Au début de chaque entretien (entre une demi-heure et une heure), il a été demandé aux enquêtés, sous forme de questions fermées, s'ils **comprenaient** (réception) chacune des langues, dialectes ou variétés en présence. Ex : Comprenez-vous le kabyle ?, Comprenez-vous l'algérien ?, etc. Dans un second temps, il a été également demandé aux enquêtés s'ils **parlaient** (production) chacune des langues, dialectes ou variétés en présence : Ex : Parlez-vous le tarifit ?, Parlez-vous le tunisien ?, etc. Le tableau suivant rend compte des réponses faites aux deux questions (réception et production) dans le cadre de questions fermées.

En %	Comprenez-vous le			Parlez-vous le		
	non	oui	un peu	non	oui	un peu
marocain	9,5	57,1	33,3	29,5 +	47,6	22,8
algérien	1,9	85,7	12,3	7,6 +	80	12,3
tunisien	8,5	64,7	26,6	19 +	53,3	27,6
kabyle	91,4	3,8	4,7	94,2 +	3,8	1,9
tamazight	95,2	2,8	1,9	97,1 +	2,8	0
tachelhit	96,1	2,8	0,9	97,1 +	2,8	0
tarifit	88,5	8,5	2,8	89,5 +	8,5	1,9
arabe litt.	22,8	59	18	22,8 =	60,9 +	16,1
français	0	57,1	42,8	0 =	60,9 +	39

Tableau 1 - Les questions fermées : différences représentationnelles entre production et réception

Ce tableau montre les différences entre les représentations que se font les locuteurs de leurs « compétences » en production et en réception. Il y a par exemple, dans la partie haute du tableau, un chiffre toujours plus élevé dans la colonne des « non » en production (parlez-vous) [cf. +] que celui des « non » en réception (comprenez-vous), ce qui tend à confirmer assez logiquement, que les enquêtés pensent mieux comprendre que parler (différence moindre en langues berbères = nombre de locuteurs moindre). Cependant, on peut voir dans la partie basse du tableau que cette tendance n'est pas confirmée dans deux cas : l'arabe littéraire et le français. On peut observer que non seulement les chiffres de la colonne des « non » en production sont les mêmes que ceux des « non » en réception, mais qu'en outre, les chiffres de la colonne des « oui » en production sont légèrement supérieurs à ceux des « oui » en réception... Or, il se trouve que l'arabe littéraire est précisément la variété haute de la diglossie maghrébine et que le français, si l'on accepte le concept de triglossie, est également une variété haute, ou au moins, si l'on refuse le concept, la langue du pays d'accueil. Sans entrer plus loin dans l'analyse, nous pouvons donc poser qu'au minimum, ces deux langues sont l'objet de représentations différentes, et qu'il n'est donc pas illogique de retrouver cette différence de statut entre H et B dans les représentations que les locuteurs se font de la production et de la réception.

2.2. Les questions ouvertes

À la fin de chaque entretien, il a été demandé aux enquêtés, sous forme de questions ouvertes, quelles langues ils **comprenaient** (réception). Ex : Quelles sont les langues que

vous comprenez ?». Puis, dans un second temps, il leur a été demandé quelles langues ils **parlaient** (production). Ex : Quelles sont les langues que vous parlez ?». Le tableau suivant rend compte des réponses faites aux deux questions (réception et production) dans le cadre des questions ouvertes.

En %	Comprenez-vous ?		Parlez-vous ?	
	oui	un peu	oui	un peu
marocain	39 +	10,4	31,4	8,5
algérien	71,4 +	5,7	68,5	3,8
tunisien	39 +	5,7	30,4	4,7
kabyle	3,8 +	0,9	2,8	0
tamazight	3,8 +	0	2,8	0
tachelhit	1,9	0,95 +	1,9	0
tarifit	7,6 =	1,9	7,6	1,9
anglais	5,7 +	10,4	4,7	5,7
arabe litt.	8,5 =	0	8,5	0
français	84,7 =	15,2	84,7	15,2

Tableau 2 - Les questions ouvertes : différences représentationnelles entre production et réception

Dans ce tableau, on observe à nouveau les différences entre les représentations que se font les locuteurs de leurs « compétences » en production et en réception. Il y a par exemple, dans la partie haute du tableau, un chiffre quasiment toujours plus élevé dans la colonne des « oui » en réception (comprenez-vous) [cf. +] que celui des « oui » en production (parlez-vous), ce qui tend à confirmer à nouveau, assez logiquement, que les enquêtés pensent mieux comprendre que parler (différence moindre en langues berbères = nombre de locuteurs moindre). Cette tendance est d'ailleurs également confirmée par la présence de l'anglais qui apparaît ici et qui « n'échappe pas à la règle ». Le seul cas pour lequel cette tendance n'est pas confirmée est le tarifit, ce qui s'explique aisément puisque seuls les locuteurs de cette « langue » s'expriment à ce sujet. En revanche, on peut une nouvelle fois observer, dans la partie basse du tableau, que la tendance à considérer que l'on « comprend » mieux que l'on ne « parle » n'est pas toujours vérifiée, et ce, pour les deux mêmes langues que dans le tableau n°1 : l'arabe littéraire et le français, langues pour lesquelles les chiffres des « oui » sont identiques dans la colonne réception (comprenez-vous) et production (parlez-vous). Ces résultats confirment de précédentes observations (Biichlé, 2003) dans le sens où les locuteurs font une distinction assez nette entre production et réception, et ce, qu'il s'agisse de questions fermées ou de questions ouvertes ; en d'autres mots, on comprend mieux que l'on ne parle. Cependant, deux langues semblent faire exception à cette « règle » : l'arabe littéraire et le français.

3. Le contraste entre réalité des pratiques et représentations

Il apparaît donc, que l'arabe littéraire et le français sont l'objet de représentations différentes des autres langues, dialectes ou variétés du Maghreb, et même, des langues étrangères, comme l'attestent les résultats de l'anglais. Cependant, la réalité des pratiques langagières en arabe littéraire et en français paraît différer sensiblement des représentations sur la production et la réception telles qu'exprimées dans les deux tableaux.

3.1. L'arabe littéraire

L'arabe littéraire ne s'apprend qu'en contexte institutionnalisé (à l'école) et cette langue est loin d'être partagée par la totalité de nos enquêtés et plus généralement, par la population maghrébine. Les taux d'analphabétisme au Maghreb étaient, en effet, de 48% au Maroc en 1999, 26,5% en Algérie en 2002, 27% en Tunisie en 1999 (sites officiels nationaux). Or,

seuls 22,8% des enquêtés déclarent ne pas comprendre l'arabe littéraire alors que 30% d'entre eux n'ont pas été scolarisés. Ce dernier point est d'ailleurs corroboré par des déclarations sur le lien entre scolarité et connaissance de l'arabe littéraire :

Ex : « *La vraie langue arabe, moi j'ai jamais fait à l'école, ni français ni arabe* » ; « *Un peu parce que je j'ai oublié, je peux prendre (apprendre) lire... et ben comme ça parler arabe ou fran arabe di /fɔsha/ (littéraire), je peux pas* ».

En outre, l'arabe littéraire n'est jamais utilisé dans la vie courante, il semble donc peu probable de bénéficier des mêmes « compétences » en production et en réception.

Ex : « *C'est difficile, il y a beaucoup de difficultés pour parler ça* ».

Nous pouvons donc formuler plusieurs hypothèses pour expliquer cette divergence entre pratiques et représentations. Comme précédemment évoqué, le prestige de l'arabe littéraire, souvent qualifié d'« arabe vrai » par les enquêtés, est renforcé par le fait que celui-ci est la langue du Coran (sourate XXVI, verset 195), plus prestigieuse, considérée comme plus belle, plus logique, plus à même d'exprimer de hautes pensées, etc. Ne pas connaître celui-ci peut donc être mal vécu pour toute personne croyante. Les personnes se trouvent alors, pour la plupart, dans cette « double contrainte » d'être, en tant qu'Arabe et musulman, censées connaître l'arabe littéraire (Grandguillaume, 1979 : 6), lequel n'est pas leur langue primaire de socialisation et qu'elles n'ont pas forcément appris à l'école. Donc, déclarer ne pas comprendre ou parler l'arabe littéraire peut être vécu comme un manque de loyauté envers sa religion ou son identité, nous évoquons ici l'arabité. Il est d'ailleurs possible de relever dans le corpus d'entretiens de nombreuses déclarations qui vont dans ce sens :

- sur la supposée normalité de connaître l'arabe littéraire : « *Dans notre religion normalement, on parle l'arabe que on comprend* » ;
- sur le lien avec la religion : « *Non...[...] un petit peu... ouais, parce que je, je fais ma prière, j'ai comprendre* » ;
- sur le prestige : « *Pour moi, la plus dure (langue), c'est l'arabe, la vraie arabe* » ;
- sur l'identité : « *Y'a des moments que tu parles avec les vrais arabes littéraires, on comprend rien [...] eux y parlent, c'est vraiment le vrai arabe, les vrais arabes, c'est pas on dit que c'est pas des vrais, c'est des vrais et nous...* »

3.2. Le français

On peut observer le même décalage entre pratiques et représentations telles qu'exprimées dans les tableaux, pour le français. En effet, pendant le déroulement de l'entretien, les déclarations sont plus nuancées et attestent la plupart du temps d'une réception supérieure à la production :

- « *Je comprends un petit peu... je peux pas répondre* » ;
- « *Je un peu je comprends mais je parler, [...] c'est difficile* » ;
- « *Chaque fois même compris tout, j'arrive pas à parler, je sais pas* ».

De plus, notre perception des « compétences » des locuteurs durant ces entretiens va également dans ce sens. Plusieurs hypothèses peuvent être formulées pour expliquer cette divergence entre pratiques et représentations. En effet, le français est la « langue » de l'ancien colonisateur, il est toujours présent dans l'administration et requis pour toutes les études supérieures. Il bénéficie donc d'un prestige certain, et de plus, dans le cas de nos enquêtés, il est la langue du pays d'accueil et symbolise un certain idéal linguistique (« parler comme il faut ! ») et au-delà, un enjeu en terme d'intégration à la nouvelle société. Beaucoup de déclarations des enquêtés vont d'ailleurs dans ce sens :

- sur la nécessité de connaître le français : « *Ici, on vit en France, on parle français* » ; « *Le français c'est bien le français, y'a pas de langue français y'a rien* » ;

- sur la relation langue/intégration : « *Heu d'apprendre le français et d'intégrer en France* » ;
- sur la nécessité de « bien » parler : « *Moi je fais toutes les possibles pour il faut bien parler* » ; « *Premièrement on parle bien français, il faut voilà... et surtout surtout pour la parole quand y parle [...] y faut bien parler* ».

4. Conclusion

Il est possible de faire plusieurs critiques à ces résultats, notamment :

- le manque de significativité statistique de certains écarts. Les écarts entre les résultats des questions sur la production et sur la réception, sont parfois ténus, ce qui peut s'expliquer par le nombre de locuteurs de certaines langues (les langues berbères) qui ne réussissent pas forcément les tests de significativité (notamment le test du χ^2). Nous nous sommes donc fondé, pour pallier cette éventuelle lacune, sur une déclaration de Labov (1976 : 305-306) à propos de ses travaux sur l'effacement du [t] et du [d] dans la communauté noire (VNA), qui estime dans une situation analogue qu'« une telle convergence de tant d'évènements indépendants suffit à nous porter à un degré de sûreté que la plupart des recherches sociales ou psychologiques continuent d'ignorer » ;
- le mode de questionnement et le profil de l'intervieweur (français) qui peuvent avoir induit des biais ;
- la prise en compte des concepts de di- ou triglossie alors que les enquêtés vivent en France. Toutefois, Lüdi & Py (1986 : 26) considèrent que dans une situation analogue (groupe de migrants dans lequel la langue d'origine est abondamment parlée), un rapprochement pourrait être fait avec « de la diglossie avec bilinguisme ».

En revanche, le fait de retrouver les mêmes tendances, pour les mêmes langues, dans les deux modes distincts de questionnement (ouvert/fermé), ne nous paraît pas fortuit. Ces contradictions entre réalité des pratiques langagières et déclarations, et entre les déclarations elles-mêmes, confirment le « poids » des représentations attachées à la « langue » et tout particulièrement, aux variétés hautes, puisque celles-ci, outre le côté purement communicationnel, sont souvent associées de manière prégnante à la religion, à l'identité, à l'intégration, à la réussite, etc. Il est d'ailleurs possible de faire le lien entre ces résultats et le fait que 60% des enquêtés, tous supposés bilingues, répondent « non » à la question « parlez-vous plusieurs langues ? ». On peut également voir, dans ces contradictions, l'expression de la double contrainte qui existe dans l'articulation entre représentations de référence (sociales et décontextualisées), et représentations en usage (individuelles et contextualisées), sachant que la tension entre ces deux entités parvient « souvent à un point de rupture proche de la contradiction formelle » (Cavalli, 2003 : 260). Nous n'avons ici, ni la volonté ni la prétention d'affirmer que ces résultats ne puissent être le fruit de quelque biais méthodologique... cependant, cette multiplication de coïncidences nous semble pour le moins source d'intérêt et constitue une manière différente de mettre en évidence la séparation fonctionnelle des variétés qui, si elle est peut-être parfois contestée dans la réalité des pratiques (moins bonne « compétence » en production, existence de continuums, de variétés intermédiaires, etc.), semble subsister au niveau représentationnel.

Références bibliographiques

ABBASSI, Abdelaziz (1977). *A sociolinguistic analysis of multilingualism in Morocco*, Thèse, University of Texas, Austin.

- ALLAL, Lamia (2004). Vivre à la croisée de deux langues : les comportements langagiers des enfants issus de l'immigration tunisienne, in *Langues et Contacts de Langues dans l'Aire méditerranéenne*, Boyer, Henri (Ed). Paris : L'Harmattan, 112-129.
- AMEUR, Meftaha & BOUMALK, Abdallah (2004). *Standardisation de l'amazighe*, Actes du séminaire organisé par le Centre de l'Aménagement Linguistique à Rabat, 8-9 décembre 2003, Publication de l'Institut Royal de la Culture Amazighe, Série : Colloques et séminaires n°3.
- BENJELLOUN, Said (2002). Une double graphie, latine et arabe, pour enseigner l'arabe marocain, in *Codification des langues de France*, Caubet, Dominique, Chaker, Salem & Sibille, Jean (Eds). Paris : L'Harmattan, 293-305.
- BIICHLE, Luc (2003). Vers un parler maghrébin de France à fonction véhiculaire et vernaculaire, Mémoire de DEA, Université Stendhal Grenoble3.
- CALVET, Louis-Jean (1999). *Pour une écologie des langues du monde*. Paris : Plon.
- CAUBET, Dominique (1998). Alternance de codes au Maghreb : Pourquoi le français est-il arabisé ?, *plurilinguismes, alternance des langues et apprentissage en contexte plurilingue, CERPL, 14* : 121-142.
- CAVALLI, Marisa (2003). *Langue, bilinguisme et représentations sociales au Val d'Aoste*. Aoste : IRRE-VDA.
- CHAKER, Salem (2002). Variation dialectale et codification graphique en berbère. Une notation usuelle pan-berbère est-elle possible ? in *Codification des langues de France*, Caubet, Dominique, Chaker, Salem & Sibille, Jean (Eds). Paris : L'Harmattan, 341-354.
- EMBARKI, Mohamed (2004). Les deux niveaux de motivation de la variation phonétique en situation de contact des langues, in *Langues et contacts de langues dans l'aire méditerranéenne*. Paris : l'Harmattan, 183-196.
- FERGUSON, Charles (1959/1972). « Diglossia », in *Language and Social Context*, Giglioli, Pier Paolo (Ed). Harmondsworth : Penguin, 232-251.
- GRANDGUILLAUME, Gilbert (1979). Langue, identité et culture nationale au Maghreb, *Peuples méditerranéens, 9* : 3-28.
- LABOV, William (1976). *Sociolinguistique*, Kihm, Alain (trad.). Paris : Minuit.
- LAROUCI, Foued (1993). Processus de minoration linguistique au Maghreb, *Cahiers de Linguistique Sociale, 22* : 45-56.
- LÜDI, Georges & PY, Bernard (1986). *Être Bilingue*. Berne : Peter Lang.
- ROMAINE, Suzanne (1989). *Bilingualism*. Oxford : Blackwell.
- TABOURET-KELLER, Andrée (1982). Entre bilinguisme et diglossie, du malaise des cloisonnements universitaires au malaise social, in *La Linguistique, revue de la société internationale de linguistique fonctionnelle*. Paris : PUF, 17-43.
- TALEB IBRAHIMI, Khoula (1995). Les Algériens et leur(s) langue(s) : éléments pour une approche sociolinguistique de la société algérienne. Alger : El Hikma.

INFLUENCE DU MILIEU SOCIAL ET DU RÉSEAU DE PAIRS SUR LES REPRÉSENTATIONS DE LA VARIATION STYLISTIQUE CHEZ DES ENFANTS PRÉADOLESCENTS FRANÇAIS

Laurence BUSON

laurence.buson@ac-grenoble.fr

Laboratoire LIDILEM - Université Stendhal Grenoble 3

38040 Grenoble Cedex 9

Abstract : This sociolinguistic study focuses on the acquisition of intra-speaker variation dealing with the way pre-teenagers interpret the stylistic variations which they are likely to perceive. Our main hypothesis is that the nature of their representations differs according to both macro-sociological factors and network factors, such as schools' heterogeneity as well as friendship dyads' heterogeneity.

1. Introduction

Si la variation stylistique (aussi appelée diaphasique ou situationnelle¹) est aujourd'hui considérée comme primordiale ou *saillante* en français (Gadet, 1998 : 62), elle est pourtant encore relativement peu étudiée, tant du point de vue de ses fonctionnements sociolinguistiques que de ses modalités d'acquisition chez l'enfant. Cet espace linguistique mal connu ouvre alors sur une nébuleuse didactique insuffisamment et inadéquatement abordée par les programmes scolaires sous l'étiquette problématique de *niveau* ou *registre* de langue².

Nous nous proposons d'étudier les représentations de la variation stylistique chez des enfants préadolescents français à travers une enquête perceptive. En effet, une meilleure connaissance de la manière dont les enfants perçoivent cette dimension de la variation et les variétés avec lesquelles ils sont en contact devrait permettre d'approcher la problématique de l'acquisition de la *souplesse* stylistique. Notons que celle-ci correspond conjointement à la conscience que possède le locuteur de l'éventail stylistique à sa disposition et à sa capacité effective à moduler son discours en fonction des caractéristiques de l'interaction (situation de communication, familiarité avec l'interlocuteur, statut de l'interlocuteur/du locuteur, intention communicative du locuteur, etc.) pour à la fois s'adapter et être acteur dans la co-construction de l'échange.

Une question, adjacente à celle de l'acquisition, concerne la part de liberté que l'on peut reconnaître au locuteur en ce qui concerne cette souplesse : dans quelle mesure est-il libre de ses choix stylistiques ou à l'inverse contraint/conditionné par des paramètres sociaux ?

Aussi nous sommes-nous intéressée à cette dimension sociale à travers la prise en compte de facteurs macrosociologiques tels que le milieu social (désormais noté *CSP* pour *catégorie socio-professionnelle*), le sexe et l'âge, mais aussi à travers des facteurs microsociologiques liés au réseau de pairs des enfants. En effet, si l'hypothèse d'une influence de la CSP doit être formulée, elle doit s'accompagner d'une réflexion plus nuancée qui s'articule autour de la

¹ Nous utiliserons ici le terme *stylistique*, pour son caractère à la fois relativement explicite et international. Il existe des fluctuations terminologiques importantes pour caractériser cette dimension de variation ainsi que les variétés qui lui sont associées, sachant que le concept même de *variété* stylistique mériterait d'être interrogé (voir Gadet, 2003, 2005).

² Notons que la question se pose également pour l'acquisition en L2, voir Tyne (2004).

mixité sociale des relations entre pairs, afin d'éviter l'écueil d'une interprétation trop déterministe qui nous semble à la fois insuffisante au plan scientifique et improductive en ce qui concerne le potentiel de remédiation.

Notre étude vise donc essentiellement la validation de deux hypothèses :

- les enfants, dès 9 ans³, sont capables d'exprimer un jugement sur des variétés stylistiques et d'explicitier leurs critères ;
- les discours des enfants sont susceptibles de varier en fonction de critères macrosociologiques, mais aussi en fonction de critères microsociologiques liés au réseau de pairs tels que la mixité sociale de l'école fréquentée ou celle des relations amicales privilégiées.

2. Méthodologie de l'enquête

Suite à une enquête préliminaire menée auprès de 36 sujets (Buson, à paraître), nous avons cherché à donner une dimension davantage quantitative à notre recherche en réalisant une deuxième étude, suivant un nouveau protocole, auprès de 196 enfants âgés de 9 à 11 ans répartis dans 6 écoles et 11 classes de l'agglomération grenobloise. Les sujets choisis fréquentent les deux dernières années de l'école primaire française, soit le CM1 et le CM2.

2.1. Echantillon et variables

Notre échantillon est composé de 107 filles et de 89 garçons issus de CSP variées, comme l'illustre le tableau ci-dessous :

	Défavorisée	Intermédiaire-	Intermédiaire+	Favorisée	
9 ans	8F + 6 G (14)	9F + 4G (13)	4F + 6G (10)	11F + 5G (16)	53
10 ans	21F + 9G (30)	13F + 8G (21)	9F + 17G (26)	9F + 11G (20)	97
11 ans	8F + 6G (14)	6F + 5G (11)	5F + 4G (9)	4F + 8G (12)	46
	58	45	45	48	196

Tableau 1 - Composition de l'échantillon

Les quatre CSP ont été déterminées à partir des informations fournies par les enseignants (fiches de renseignements remplies par les parents eux-mêmes, ceux-ci indiquant leurs professions) et par les enfants (lors d'un entretien individuel, ils étaient invités à nous préciser le métier de leurs parents). Un indice social, inspiré des catégories de l'INSEE, a ainsi été établi à partir de la profession des deux parents et d'éventuels beaux-parents en cas de garde alternée.

Pour aborder la question de la mixité sociale du réseau de pairs, nous avons retenu deux variables : la mixité sociale des dyades d'amis et la mixité sociale de l'école. Nous avons pris en compte les choix amicaux déclarés des enfants : chacun a été amené à désigner son meilleur ami dans la classe⁴ et la mixité de la dyade a ainsi pu être déterminée par la combinaison des milieux sociaux de chacun. Ainsi, nous considérons que l'enfant appartient à une dyade à mixité sociale forte s'il existe deux niveaux entre les CSP de la dyade (ex :

³ Notons que Labov (1976 : 207), dans la mesure où il s'intéresse à une maîtrise totale et non à une compétence en cours de construction, situe l'acquisition de la variation stylistique bien plus tard, à la fin de l'adolescence.

⁴ L'enfant devait désigner ses meilleurs amis et ordonner ses choix. La dyade préférentielle a été définie à partir de l'enfant cité en 1^{er} choix. Si ce dernier n'était pas identifié dans l'échantillon, le 2^{ème} puis le 3^{ème} choix ont été pris en compte. Si aucun "meilleur ami" ne pouvait être identifié, l'enfant a été catégorisé "hors dyade" et n'a pas été pris en compte dans l'analyse pour cette variable. C'est le cas de 10 sujets sur le total de l'échantillon.

défavorisé/favorisé ou intermédiaire+/défavorisé⁵). Dans le cas contraire, la mixité est dite faible ou nulle (ex : fav/fav ou inter-/défav). Nous notons alors que 41 enfants relèvent de dyades mixtes et 145 de dyades peu ou non mixtes.

Concernant le profil social de l'établissement, nous considérons qu'une école peut être qualifiée de "non mixte" dans la mesure où plus de 80% des enfants qui la fréquentent appartiennent au même milieu social. Dans notre échantillon, 26 enfants appartiennent à une école non mixte⁶.

2.2. Protocole de recueil de données

Les entretiens individuels de 15/20 minutes se sont déroulés comme suit. L'enfant :

- écoute un enregistrement de trois énoncés variant du point de vue du degré de formalité (3 messages d'accueil de répondeur),
- est invité à développer un commentaire non dirigé initié par une question ouverte avec relance ("qu'est-ce que tu en penses ?"),
- répond à un questionnaire concernant les énoncés entendus⁷,
- est invité à donner quelques renseignements personnels sur lui et ses relations sociales⁸ (questions standardisées sur sa famille, son réseau de pairs, ses habitudes en terme de lecture, etc.),
- se prête à la mise en scène d'un échange fictif très bref avec trois interlocuteurs plus ou moins familiers (interpeller sa maman, un copain, un inconnu au sujet d'un blouson perdu).

Le questionnaire concerne l'énoncé le plus formel puis le plus informel et s'articule autour de trois grands points :

- un jugement sur la variété entendue (ça peut m'arriver ou non de parler comme ça, j'aime/j'aime pas sa façon de parler, etc.),
- la perception que l'on peut avoir du locuteur en l'écoutant (l'enfant doit choisir ou rejeter des étiquettes correspondant à différents éclairages sur le locuteur, avant d'ordonner celles qu'il a gardées ; les propositions concernent son milieu social, sa capacité à être poli ou non, son intention communicative, l'attention portée au discours, etc.),
- enfin une évaluation de l'interlocuteur supposé, en terme de familiarité et de statut (on parle comme ça à quelqu'un qu'on connaît bien/qu'on ne connaît pas, quelqu'un d'important/un ami, etc.).

2.3. Protocole d'analyse des données

Les commentaires spontanés obtenus en réaction à l'écoute de l'enregistrement feront l'objet d'une analyse qualitative ultérieure ; dans cet article, nous présenterons une analyse quantitative des données. Deux indices ont été constitués à partir de scores attribués aux différents types de réponses, de manière à prendre en compte un ensemble d'éléments convergents :

- l'indice de référence à la norme (Rn)
- l'indice de référence à l'interaction (Ri)

⁵ Désormais, les quatre CSP seront notées comme suit : "défav", "inter-", "inter+" et "fav".

⁶ Il s'agit d'une école défavorisée.

⁷ Les thèmes abordés, catégories et formulations présents dans le questionnaire sont issus de l'enquête préliminaire et correspondent à des réponses données par les enfants.

⁸ Précisons que nous nous sommes inspirée ici, de manière assez libre et non exhaustive, des méthodes sociométriques connues en psychologie sociale.

Ces indices combinent les réponses issues de la phase d'entretien et du questionnaire. Rn est élevé si l'enfant s'appuie sur des critères normatifs pour interpréter la variation entendue : il émet des jugements de valeur, fait référence à la politesse, au bien et au mal parler, etc. Ri est élevé si l'enfant fait référence aux caractéristiques de l'interaction pour interpréter la variation entendue : il privilégie un locuteur libre à un locuteur contraint et s'appuie par exemple sur le critère de la familiarité avec l'interlocuteur, sur l'intention communicative du locuteur, etc.

Il serait trop fastidieux d'entrer ici dans le détail de l'attribution des scores. Nous proposons néanmoins ci-dessous, pour tenter de rendre plus explicite le fonctionnement de ces indices, des exemples de réponses pour Ri :

a. Question ouverte non dirigée :

"c'est un monsieur qui fait plusieurs répondeurs - pour ses copains - et pour les autres --- parce que quand même 'salut les gars' ça se dit pas à tout le monde -- les autres c'est plutôt par exemple ceux qui appellent au boulot - par exemple son chef de service je sais pas moi -- la 1ère elle est plutôt bien /.../ moi si c'est quelqu'un que j'aime pas beaucoup je vais quand même lui parler un peu gentiment mais bon pas avec des fleurs dans les bulles -- ça pourrait m'arriver par exemple si c'est quelqu'un d'assez important qui appelle bah je vais parler poliment".

Notons que chaque commentaire non dirigé a été analysé de manière à rechercher des références à l'interaction et/ou à la norme. Afin d'évaluer le caractère aléatoire de cette analyse, nous avons procédé à un test de stabilité⁹ pour 103 sujets sur les 196 ; le pourcentage de stabilité obtenu est de 93%.

Dans la déclaration ci-dessus, on considère que l'enfant fait référence à la politesse ainsi qu'à la relation à l'interlocuteur, à travers les notions de statut et d'affinité : des points sont donc accordés pour le score de Rn comme pour celui de Ri (pour Ri, voir les passages soulignés). Cet exemple illustre par ailleurs que Ri et Rn ne sont pas inversement proportionnels : l'acquisition de la souplesse stylistique chez l'enfant prend appui à la fois sur les conventions sociales et sur la prise en compte des paramètres de l'interaction, sans que cela soit contradictoire.

b. Questionnaire 1^{ère} partie portant sur le thème "attirance/rejet d'une variété" :

Le choix des modalités formulées positivement ("ça peut m'arriver de parler comme ça", "je connais des gens qui parlent comme ça", "j'aime bien comme il parle") augmentera l'indice de Ri.

c. Questionnaire 2^{ème} partie portant sur le thème "se représenter le locuteur" :

Le rejet des modalités où le locuteur est perçu comme contraint par des paramètres externes (rejet de "il est riche" & "il est pauvre"), combiné au choix des modalités concernant l'attention portée au discours et l'intention communicative du locuteur ("il fait attention à sa façon de parler", "il veut avoir l'air sérieux"), augmentera l'indice de Ri.

d. Questionnaire 3^{ème} partie portant sur le thème "se représenter l'interlocuteur" :

Le choix des modalités proposant un interlocuteur non familier et un interlocuteur à statut élevé pour l'énoncé formel¹⁰ augmentera l'indice de Ri.

⁹ Accord intra-juge : codage effectué à deux reprises, à cinq mois d'intervalle.

¹⁰ ...et vice versa pour l'énoncé informel : pour les parties 1, 2 et 3 du questionnaire, les mêmes questions (ou leur pendant) étaient posées deux fois, pour l'énoncé le plus formel et pour l'énoncé le plus informel.

e. Simulation d'échange :

La présence de modalisations en situation formelle et pas en situation informelle augmentera l'indice de Ri. Le degré de modalisation correspond à plusieurs critères, dont la présence de termes d'adresse, de salutation et d'excuse ainsi que l'emploi du conditionnel.

(à un ami) "j'ai perdu mon blouson est-c(e) que tu l'as pas vu?"

(à un inconnu) "excusez-moi d(e) vous déranger monsieur? est-c(e) que vous aur(i)ez vu un blouson?"

L'analyse des résultats abordera en premier lieu l'influence des variables macrosociologiques (sexe, âge et CSP) pour chacun des deux indices, puis les variables microsociologiques pour l'indice Ri.

3. A l'influence de la CSP, s'ajoute celle de la mixité sociale des relations entre pairs

Pour les variables "macro", nous avons procédé à une analyse de variance (ANOVA).

3.1. L'influence de la CSP et de l'âge sur l'indice de référence à la norme (Rn)

Nous avons cherché à observer les effets du sexe, de l'âge et de la CSP sur Rn. Comme l'illustre la figure 1¹¹, deux phénomènes sont à remarquer : une plus forte normativité chez les enfants de CSP- d'une part et une convergence des réponses à partir de 10 ans d'autre part¹². L'ANOVA révèle un effet significatif de la CSP ($F_{1,188} = 5,88$; $p = 0,01$) ainsi qu'une interaction entre âge et CSP si l'on oppose les groupes d'âge 9 ans et 10-11 ans ($F_{1,188} = 5,18$; $p = 0,02$). En effet, les enfants de CSP- et en particulier les inter- se révèlent particulièrement normatifs, et ce plus spécifiquement avant 10 ans.

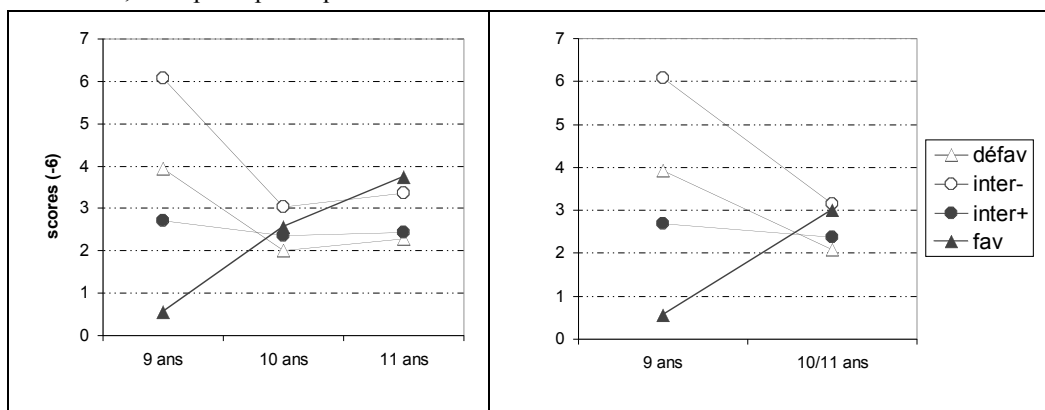


Figure 1 & 1bis - Mise en évidence de l'influence de l'âge et de la CSP sur Rn

Ce constat n'est pas particulièrement surprenant dans la mesure où l'on observe la même tendance à l'hypernormativité des classes intermédiaires chez l'adulte. En effet, de nombreuses recherches (voir par exemple Labov, 1976 ou Bourdieu, 1982) sous-tendent l'existence du poids des normes variant de manière non linéaire entre les CSP, comme l'explique Perrenoud (1988 : 31) :

¹¹ L'ANOVA a été réalisée avec un regroupement en deux CSP (moins et plus) pour conserver un nombre suffisant de sujets ; pour plus de lisibilité, les figures détaillent les quatre CSP.

¹² Cette tendance s'observe pour les deux sexes.

"Dans certains milieux privilégiés, on ne déteste pas un certain relâchement, l'affectation d'un parler populaire ou une certaine familiarité, signes de décontraction et d'aisance. C'est généralement dans la fraction la plus conservatrice des classes moyennes qu'on observe la plus forte propension à prendre la norme scolaire au sérieux, voire au tragique."

Néanmoins, alors qu'à 9 ans, les écarts sont significatifs entre les CSP¹³, on observe une baisse significative de la référence à la norme chez les inter- (PLSD de Fisher, $p = 0,04$) qui aboutit à une situation de convergence entre les CSP. L'hypothèse interprétative que nous privilégions ici est celle du caractère immature de la sur-référence à la norme, susceptible de s'effacer à l'adolescence. En effet, si l'enfant de 9 ans, encore très empreint du discours parental, se fait le reflet de la prégnance des conventions normatives de son milieu, on peut supposer qu'en grandissant, il privilégie la relation entre pairs (Bauvois, 1998). Les jugements ont alors tendance à s'homogénéiser et il se détache ainsi de ces références ascendantes saillantes, en particulier si cette référence est une sur-référence comme dans le cas des enfants inter-. Les interprétations normatives de la variation restent donc présentes chez tous les enfants, de manière assez équivalente, sans être centrales ni discriminantes.

Si la norme est un critère omniprésent dans le discours sur la variation stylistique, il ne s'agit pas selon nous du critère le plus prégnant dans la mesure où il laisse de côté la part de liberté du locuteur, en soumettant ce dernier à des codes qu'il se doit d'intérioriser et de respecter et sur lesquels il n'a pas la possibilité d'agir. Cette vision dichotomique de la variation, où l'on parle bien ou mal "dans l'absolu", est mise en perspective et nuancée par les références à l'interaction qui, elles, contextualisent le discours pour approcher l'idée de parler bien ou mal "relativement" à une situation de communication qui, de plus, n'est pas donnée *a priori* mais à construire entre les interlocuteurs. C'est pourquoi l'indice Ri représente pour nous l'indice premier de construction de la *souplesse* stylistique.

3.2. La mixité sociale du réseau de pairs vient nuancer l'influence de la CSP sur l'indice de référence à l'interaction (Ri)

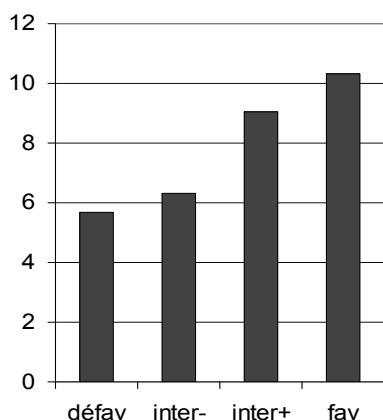


Figure 2 - Influence de la CSP sur Ri

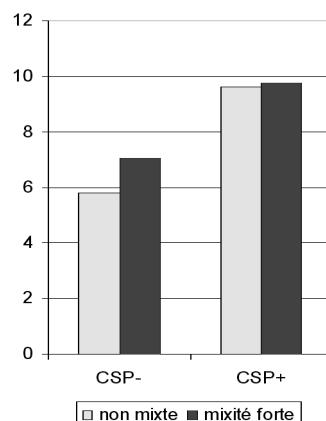


Figure 3 - Influence de la mixité de la dyade sur Ri

¹³ D'après le test *a posteriori* LSD de Fisher réalisé avec les quatre CSP, à 9 ans, les écarts sont significatifs entre "fav" et "défav" ($p = 0,03$) et entre "fav" et "inter-" ($p = 0,001$).

L'ANOVA à trois facteurs (CSP, âge, sexe) révèle un effet significatif de la CSP sur Ri ($F_{1, 184} = 22,29$; $p < 0,0001$)¹. Comme l'illustre la figure 2, plus le milieu social est favorisé, plus la tendance à faire référence à l'interaction est forte.

Néanmoins, ce constat est susceptible d'être nuancé par la variable de mixité sociale des dyades d'amis.

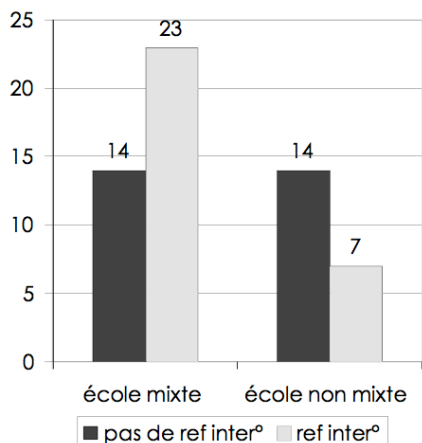


Figure 4 - Influence de la mixité sociale de l'école sur la référence spontanée à l'interaction (question ouverte non dirigée) chez les enfants issus de milieux défavorisés

En effet, si une mixité forte ne renverse pas la tendance selon laquelle les scores des CSP- sont inférieurs à ceux des CSP+, on observe tout de même une réduction nette de cet écart, réduction qui correspond à des scores plus élevés chez les CSP- (cf. figure 3). Ainsi, l'écart entre CSP, qui est significatif dans le cas des dyades non mixtes ($F_{1,143} = 17,24$; $p < 0,0001$), ne l'est plus quand les dyades sont mixtes ($F_{1,39} = 1,73$; $p = 0,19$).

En outre, les analyses plus qualitatives des entretiens révèlent des tendances similaires.

Ainsi, comme l'illustre la figure 4 ci-contre, si nous comparons les réponses à la question ouverte des enfants issus de milieux défavorisés fréquentant des écoles socialement mixtes et ceux fréquentant des écoles non mixtes, nous constatons que les enfants faisant spontanément référence à l'interaction sont significativement plus nombreux dans les écoles mixtes ($\chi^2 = 4,458$ à 1ddl ; $p = 0,034$). Ces observations

corroborent donc les résultats globaux que nous avons obtenus pour Ri.

C'est pourquoi nous formulons l'hypothèse selon laquelle le fait d'avoir des relations sociales nombreuses, riches et diversifiées favoriserait le développement de représentations ouvertes, privilégiant la dynamique interactionnelle à une image sociale figée du locuteur. En effet, le locuteur peut alors être perçu comme un acteur qui a la possibilité de s'adapter à l'interlocuteur et au contexte de l'interaction, et d'agir dans la co-construction de celle-ci, ce qui correspond à la manière dont nous envisageons la *souplesse* stylistique.

4. Conclusion

Nous observons donc une certaine convergence au sein des résultats de cette enquête. Même si l'on se doit bien sûr de rester prudent par rapport aux biais susceptibles d'être induits par la méthode du questionnaire en contexte scolaire (comme l'explique Jeanjean (1978 : 65), cette méthode comporte toujours le risque de mesurer davantage "une certaine aptitude à se plier aux consignes, une certaine conformité au modèle du 'bon élève', en un mot un certain conformisme."), la CSP apparaît comme une variable incontournable pour l'analyse des représentations de la variation stylistique. Néanmoins, il est sans doute intéressant de mettre également en lumière l'influence des variables de réseau pour une analyse plus fine des situations. La mixité sociale¹⁴ du réseau de pairs semble avoir une influence non négligeable sur les représentations des enfants en ce qui concerne les commentaires sur les variétés stylistiques et les critères invoqués. Après ce travail sur les représentations, reste à étudier les pratiques langagières effectives des enfants et à les confronter à ces mêmes variables, afin d'appréhender leur éventuelle influence sur la variété et l'étendue des répertoires stylistiques.

¹⁴ Voir Varro (2003) pour une revue détaillée du concept de mixité en sociologie.

En définitive, si le milieu social, interagissant ou non avec l'âge, semble avoir un effet sur la nature des critères invoqués, il peut être intéressant d'explorer des voies moins pré-déterminées sur lesquelles les individus ou les institutions peuvent éventuellement agir. On constate à ce propos que l'actualité nous présente souvent la problématique de la mixité sociale et scolaire comme une évidence salvatrice *a priori* ; il est en revanche beaucoup plus rare d'avoir accès à des études rigoureuses mettant en perspective mixité sociale et réussite scolaire.

Références bibliographiques

- BAUVOIS, Cécile (1998). L'âge de la parole : la variable âge en sociolinguistique, in *DiversCité Langues*, Vol. III, revue en ligne, <http://www.uquebec.ca/diverscite>. Consulté en septembre 2006.
- BOURDIEU, Pierre (1982). *Ce que parler veut dire. L'économie des échanges linguistiques*, Paris : Fayard.
- BUSON, Laurence (à paraître). La variation stylistique chez les enfants de 10/11 ans : une étude exploratoire en contexte français, in *Actes du colloque d'Oxford*, juin 2005.
- GADET, Françoise (1998). Cette dimension de variation que l'on ne sait nommer, in *Sociolinguistica*, n°12, 53-71.
- GADET, Françoise (2003). *La variation sociale en français*, Paris : Ophrys.
- GADET, Françoise (2005). Research on sociolinguistic style, in *Sociolinguistics*, Ammon, U., Dittmar, N., Mattheier, K., Trudgill, P. (Eds). Berlin & New York : De Gruyter, 1353-61.
- JEANJEAN, Colette (1978). Et si le handicap était produit par l'institution scolaire ?, in *Le handicap socio-culturel en question*, C.R.E.S.A.S., E.S.F., 65-82.
- LABOV, William (1976). *Sociolinguistique*, Paris : Editions de Minuit.
- PERRENOUD, Philippe (1988). "Parle comme il faut !". Réflexions sociologiques sur l'ordre linguistique, in *La langue française est-elle gouvernable ? Normes et activités langagières*, Schoeni, G., Bronckart, J.-P. et. Perrenoud, Ph. (dir). Neuchâtel et Paris : Delachaux et Niestlé, <http://www.unige.ch>. Consulté en août 2006.
- TYNE, Henry (2004). Le style en français langue seconde : vers un regard sociolinguistique, in *Langage et Société*, n°109, 31-52.
- VARRO, Gabrielle (2003). *Sociologie de la mixité. De la mixité amoureuse aux mixités sociales et culturelles*, Paris : Belin.

L'ENVIRONNEMENT LANGAGIER DU TRÈS JEUNE ENFANT DANS LA SOCIÉTÉ PLURILINGUE DE VÉNÉTIE : DESCRIPTION D'INDICES STATISTIQUES ET PRAGMATIQUES¹

Anna GHIMENTON

anna.ghimenton@u-grenoble3.fr

LIDILEM

Université Stendhal Grenoble 3, BP25

38040 Grenoble Cedex 9

Abstract : In a multilingual society, variation in language usage is the norm. The main focus of this study is the variation in the language choices of adults in a child's language environment. We focus on the language organization of the adults' speech turns when addressing the child directly as opposed to other adults, in order to explore the importance of the language environment during language acquisition.

1. Introduction

Depuis l'unification de l'Italie, les parlers italo-romans sont encore utilisés, malgré le processus d'« italianisation » de ces variétés minoritaires en cours depuis quelques années (Sanga, 1981). De ce fait, l'espace italo-roman reste une aire plurilingue caractérisée par plusieurs zones de contact entre l'italien et les dialectes régionaux. Ces contacts sont à l'origine, d'une part, de la formation des italiens régionaux (cf. Telmon, 1990), d'autre part, des phénomènes de variation généralisée qui prennent la forme d'alternances codiques (cf. De Mauro, 1963 ; Berruto, 1995 ; Sabatini, 1997). Notre attention se focalisera sur la Vénétie, une région nord-orientale italienne, dont le dialecte a une tradition littéraire² et un prestige socio-politique importants : jusqu'en 1797, il était en effet la langue officielle de la Serenissima, l'ancienne République de Venise.

Aujourd'hui, le dialecte de Vénétie bénéficie d'une vitalité remarquable comme le montre un sondage par questionnaire mené par l'ISTAT³ en 2000. Les résultats attestent que la langue employée quotidiennement en interaction avec les membres de la parenté et avec les amis est de préférence le dialecte, tandis que l'italien est utilisé majoritairement lors des échanges avec des gens extérieurs au cercle des connaissances proches⁴. La volonté de conserver le dialecte serait « le fruit d'un choix libre » (Marcato, 1981) de la part du locuteur. L'emploi de cette langue marquerait le degré de familiarité entre les locuteurs. Cependant, malgré le lien affectif et la connotation familiale qui accompagnent l'emploi du dialecte, il est de plus en plus fréquent, comme le signale Giacalone-Ramat (1995), d'utiliser de préférence l'italien en s'adressant aux membres plus jeunes de la parenté. Un des buts de cette étude est d'explorer les raisons de cet usage plus marqué de l'italien en direction des plus jeunes.

¹ Cette étude a bénéficiée d'un financement octroyé par le Ministère délégué à la Recherche et aux Nouvelles technologies : Action Concertée Incitative, Espace et Territoire.

² Il suffit de penser à Carlo Goldoni, un dramaturge vénitien du XVIII^{ème} siècle.

³ L'Institut National Italien des Recherches Statistiques. Les résultats du sondage ont été publiés dans le quotidien national *Gazzettino*, le 16 Septembre 2002.

⁴ Ces mêmes résultats ont été trouvés auparavant dans une enquête, menée en 1998 par l'Institut *Poster* auprès de 1200 sujets interviewés. Les résultats de ce sondage ont été publiés dans le quotidien national *Gazzettino*, le 14 décembre 1998.

2. Présentation de l'étude de cas

Cette étude fait partie d'un projet plus vaste⁵ dont l'objet est d'étudier la relation entre l'environnement langagier des enfants de Vénétie et le développement plurilingue. Les enquêtes se déroulent dans la ville de Castelfranco Veneto (environ 31000 habitants) située dans la partie centrale de la Vénétie. Nous présenterons ici une étude de cas destinée à décrire les usages adressés à un enfant, Francesco, ou produits en sa présence. Il est âgé d'un an et un mois et produit essentiellement du babillage répété. Les résultats présentés sont issus de l'analyse de quatre heures d'enregistrement (soit 1702 tours de parole au total) concernant des interactions impliquant cinq membres de la parenté maternelle. Outre le père et la mère, les trois autres locuteurs sont la tante et les grands-parents de l'enfant. Chaque mot de chaque tour de parole a été codé selon son appartenance à l'une des trois catégories linguistiques suivantes : dialecte, italien et continuum. Plus précisément, nous avons attribué chaque mot au pôle « dialecte » ou au pôle « italien », et les mots appartenant aussi bien à un code qu'à l'autre ont été catégorisés en tant qu'unités lexicales situées sur le continuum linguistique. Le plurilinguisme offre une variété de choix énonciatifs qui puisent soit dans des zones de divergence, soit dans des zones de convergence des répertoires en contact ; une catégorie de langue comme celle du continuum s'avère donc nécessaire pour classer les usages composés de formes linguistiques qui appartiennent à l'italien aussi bien qu'au dialecte. Il faut préciser que notre attention ne se portera que sur les usages langagiers des adultes en présence de l'enfant, de façon à décrire l'environnement familial précoce. En effet, l'input fourni par le microcosme familial exerce, sur la construction d'un répertoire plurilingue, une influence plus grande que celle des langues présentes dans la communauté élargie (De Houwer, 1990). En observant les usages des adultes en présence de Francesco, nous essayerons d'en déduire leurs représentations en matière d'éducation langagière au sein d'une société où l'alternance codique est une pratique quotidienne comme dans de nombreux contextes plurilingues. Ce phénomène contribue à un environnement caractérisé par une forte variation des usages langagiers, comme c'est le cas dans la société de Vénétie⁶.

Notre questionnaire général porte sur les trois points suivants : 1/ Comment les adultes gèrent-ils l'alternance codique lorsqu'ils s'adressent à un enfant construisant son répertoire plutôt qu'à un autre adulte ? 2/ Quels sont les repères linguistiques, statistiques et pragmatiques qui favorisent une telle construction ? 3/ Est-ce que les usages des parents s'adressant à l'enfant laissent apparaître des visées éducatives langagières ? Pour explorer ces questions, nous commencerons par présenter une analyse générale des usages des adultes en présence de l'enfant. Ensuite, nous centrerons notre examen sur les énoncés adressés à l'enfant afin de mettre en relief d'éventuelles attitudes éducatives.

2.1. Analyse 1 : Répartition des langues dans les tours de parole

Dans cette partie de notre étude, nous examinerons la répartition des langues dans les tours de parole. Nous calculerons les pourcentages moyens de mots italiens, de mots dialectaux et de mots du continuum, produits à l'intérieur des tours de parole de chaque locuteur. Par exemple, dans le tableau ci-dessous, les énoncés de la tante de Francesco sont composés de 46.3% de mots dialectaux, de 16.8% de mots italiens et 36.8% de mots du continuum.

⁵ Ce projet s'inscrit dans le cadre d'une thèse de doctorat.

⁶ Cf. Berruto (2006).

Locuteur	DIALECTE	CONTINUUM	ITALIEN
Tante (25 ans)	45.8% (298)	36.2% (298)	17.8% (298)
Grand-mère (60 ans)	42.1% (458)	34.6% (458)	23.1% (458)
Grand-père (64 ans)	49.4% (227)	35.4% (227)	15.2% (227)
Mère (30 ans)	39.1% (434)	34.8% (434)	26.1% (434)
Père (31 ans)	34.6% (189)	35.6% (189)	29.7% (189)
Kruskall- Wallis ⁷	H = 25.9 ; p < 0.0001	H = 1.18 ; p = 0.88	H = 47.03 ; p < 0.0001

Tableau 1 – Moyennes des pourcentages de mots provenant du dialecte, du continuum et de l'italien par tour de parole ; entre parenthèses : nombre de tours ayant contribué au calcul des pourcentages moyens

L'observation des taux d'usages des langues nous permet de mettre en évidence une tendance générale : les adultes emploient globalement davantage de formes dialectales que de formes italiennes ou de formes appartenant à la catégorie du continuum. Nous avons évalué la significativité des différences entre les taux individuels d'usage obtenus pour chaque catégorie de langue ; ces valeurs sont toutes significativement différentes. La production dialectale des adultes s'échelonne entre 34.6% pour le père et 49.4% pour le grand-père. Les taux des usages qui se situent sur le continuum se répartissent dans une fourchette réduite : entre 35.4% produit par la grand-mère à 36.2% réalisé par la tante. Enfin, les moyennes d'usage de l'italien sont les plus basses : elles s'étalent entre 15.2% produit par le grand-père et 29.7% réalisé par le père.

Si maintenant nous nous intéressons aux différences entre locuteurs, nous constatons que les parents et la grand-mère de l'enfant utilisent moins de dialecte que la tante ou le grand-père. Qu'il s'agisse de la production de l'italien ou du dialecte, les parents de Francesco se situent d'ailleurs systématiquement aux extrêmes de l'échelle : ils sont les deux locuteurs dont la fréquence d'usage de l'italien serait la plus élevée et la fréquence d'usage dialectal la plus faible. A ce stade, une question se pose : cet usage de l'italien plus marqué chez les parents traduit-il une visée éducative ? Dès les premiers mois de vie de l'enfant, les parents favoriseraient l'italien plutôt que le dialecte lorsqu'ils s'adressent à celui-ci. Ainsi, les scores plus faibles de production dialectale de la part des parents pourraient être dus à l'évitement du dialecte dans les énoncés qu'ils adressent à l'enfant.

Pour donner suite à ce questionnaire, nous avons procédé de la façon suivante. D'abord, nous avons classé chaque tour de parole en fonction de la distance énonciative par rapport à l'enfant. Plus précisément, les tours de parole ont été répartis entre les trois catégories suivantes : 1/ les tours produits dans une interaction entre adultes où l'on fait référence à l'enfant ; 2/ les tours produits lors d'une conversation entre adultes dont le contenu ne porte pas sur l'enfant ; 3/ les tours produits dans des énoncés adressés à l'enfant.

2.2. Analyse 2 : Usages des langues en fonction de l'interlocuteur

Dans un premier temps, nous avons regroupé les deux premiers types de tours de parole⁸ en les opposant au troisième. D'un côté, nous retrouvons les tours où les adultes parlent entre

⁷ Le test de Kruskal Wallis est un test statistique non paramétrique qui détermine si plus de deux moyennes calculées sur des séries indépendantes peuvent être considérées comme significativement différentes.

⁸ Voir les trois lignes directrices lors de la répartition des tours de parole.

eux et de l’autre côté, les tours que les adultes adressent à l’enfant. Pour chaque locuteur et chaque type de tour, nous avons calculé les pourcentages moyens de mots italiens et de mots dialectaux. Etant donnée l’homogénéité interindividuelle des usages qui se situent sur le continuum linguistique (voir analyse 2.1) cette dernière catégorie ne sera plus prise en compte. Notre attention se portera donc sur les unités lexicales qui appartiennent clairement aux systèmes de l’italien et du dialecte et non à la zone indifférenciée du continuum.

L’hypothèse émise pour explorer les variations dans les usages adultes est la suivante : est-ce que le statut de l’interlocuteur - enfant *versus* autre adulte - entraîne une différenciation du choix entre italien et dialecte ?

	Dialecte vers adultes	Dialecte vers enfant	Italien vers adultes	Italien vers enfant
Tante (25 ans)	50.9% (267)	2.6% (31)	11.8% (267)	69.5% (31)
	U = 762 ; p < 0.001		U = 999.5 ; p < 0.0001	
Grand-mère (60 ans)	50% (359)	13.4 (99)	14.5% (359)	54.4% (99)
	U = 6317.5 ; p < 0.0001		U = 5707 ; p < 0.0001	
Grand-père (64 ans)	51% (202)	35.8% (25)	12% (202)	40.7% (25)
	U = 1887.5 ; p < 0.03		U = 1533 ; p < 0.0002	
Mère (30 ans)	52.9% (313)	3.3% (121)	9.7% (313)	68.7% (121)
	U = 3412 ; p < 0.0001		U = 2098 ; p < 0.0001	
Père (31 ans)	50% (128)	2.4% (61)	12.9% (128)	65% (61)
	U = 563 ; p < 0.0001		U = 805 ; p < 0.0001	

Tableau 2 – Moyennes des pourcentages de mots dialectaux et italiens par tour de parole ; entre parenthèses : le nombre de tours ayant contribué au calcul des pourcentages moyens ; en italiques : test de Mann-Whitney

Le tableau ci-dessus représente les pourcentages⁹ moyens par individu de mots dialectaux et de mots italiens à l’intérieur des tours de parole. Les productions dialectales des adultes en direction d’un autre adulte varient entre 50% (produit par le père et la grand-mère) et 52.9% (produit par la mère). Au contraire, le dialecte est moins présent dans les tours adressés à l’enfant. Dans ce type de tour, tous les adultes diminuent significativement¹⁰ leur production dialectale, qui s’échelonne alors entre 2.4% (produit par le père) et 35.8% (réalisé par le grand-père).

Par ailleurs, les usages de l’italien manifestent une tendance inverse par rapport aux usages dialectaux décrits précédemment. En direction des autres adultes, l’italien est la langue la moins utilisée, avec un éventail de production restreint qui s’étale entre 9.7% (produit par la mère) et 14.5% (produit par la grand-mère). En revanche, en direction de l’enfant, tous les locuteurs augmentent les taux d’italien de manière statistiquement significative. Les pourcentages minimum et maximum de mots italiens par tour de parole sont respectivement de 40.7% (grand-père) et 69.5% (tante).

Une première conclusion s’impose : tous les locuteurs modulent significativement leur production de l’italien et du dialecte selon le niveau de développement langagier de son interlocuteur. Lorsqu’ils sont adressés à un adulte, les tours de parole des locuteurs sont

⁹ Pour rappel, nous n’avons pas inclus dans le tableau les taux d’usage de mots provenant de la catégorie du continuum. Donc, la somme des pourcentages figurant dans le tableau n’atteint pas 100%.

¹⁰ Pour ce faire, nous avons utilisé le test non paramétrique de Mann Whitney qui établit si la différence entre deux moyennes est statistiquement significative (au taux de p = 0.05).

composés majoritairement de mots dialectaux. À l'inverse, les mots qui composent les tours de parole adressés à l'enfant, appartiennent majoritairement au lexique italien.

Par ailleurs, on constate que cette tendance à limiter le dialecte et à préférer l'italien pour parler à l'enfant est plus ou moins marquée selon l'âge du locuteur. S'adressant à l'enfant, le père, la mère et la tante limitent leur emploi du dialecte : leurs usages de cette langue se répartissent entre 2.4% et 3.3%. Cependant, la grand-mère et le grand-père ont une production dialectale plus importante avec respectivement 13.4% et 35.8% de dialecte. La tendance inverse se manifeste à l'égard de l'italien en direction de l'enfant : le grand-père et la grand-mère produisent moins de mots italiens que les trois autres locuteurs.

Les résultats obtenus corroborent et précisent notre hypothèse de travail initiale. Même si tous les locuteurs augmentent la proportion de mots italiens et diminuent la part des mots dialectaux lors d'un échange avec l'enfant, tous ne le font pas dans la même proportion. Cette tendance est en effet liée à la variable « âge », la préférence de l'italien au détriment du dialecte dans les énoncés vers l'enfant étant plus marquée chez les adultes plus jeunes, et tout spécifiquement les parents de l'enfant.

Une interprétation directe de la prévalence de l'italien dans les usages des jeunes adultes en direction de l'enfant consiste à invoquer des motivations éducatives. Cependant, si nous regardons les taux de dialecte et d'italien produits avec les adultes, nous remarquons que, même si les différences entre locuteurs dans chacune de ces catégories sont relativement faibles, certaines se révèlent significatives, comme nous le verrons ci-dessous. Comme le montre le test de Kruskal Wallis ($H = 20,086$; $p = 0,0005$) tous les adultes ne produisent pas la même proportion de mots italiens quand ils s'adressent à un autre adulte. Il se peut que les préoccupations éducatives des parents à l'égard de l'enfant se manifestent même au travers des interactions avec les autres adultes en présence de l'enfant.

2.3. Analyse 3 : Usages de l'italien dans les échanges entre adultes

Pour explorer l'hypothèse concernant les préoccupations éducatives des parents, nous nous focaliserons sur la répartition des mots italiens à l'intérieur des échanges entre adultes. Nous avons donc réparti les tours de parole des adultes en deux catégories, selon les critères thématiques suivants : 1/ les tours dans une interaction entre adultes dont le contenu porte sur l'enfant 2/ les tours produits lors d'une conversation entre adultes où l'enfant n'est pas évoqué. Tous les échanges dont il est question se sont déroulés en présence de l'enfant.

ITALIEN PRODUIT DANS UNE INTERACTION ENTRE ADULTES			Mann Whitney
Locuteur	Thème : Autre	Thème : Enfant	
Tante	11.9% (224)	10.9% (43)	$U = 4607$; $p = 0.6051$
Grand-mère	16% (262)	10.2% (97)	$U = 12008$; $p = 0.3912$
Grand-père	13.3% (153)	8.2% (49)	$U = 3460.5$; $p = 0.4187$
Mère	10.4% (225)	7.8% (88)	$U = 9571$; $p = 0.5978$
Père	11.6% (102)	17.8% (26)	$U = 1080$; $p = 0.10$

Tableau 3 – Moyennes des pourcentages de mots italiens par thème abordé ; entre parenthèses : le nombre de tours ayant contribué au calcul des pourcentages moyens ; en italiques : test de Mann-Whitney

La baisse apparente de la proportion de mots italiens pour la tante, le grand-père, la grand-mère et la mère n'est ni significative, ni tendancielle (Mann-Whitney $> 0,10$). En revanche, les usages du père montrent une augmentation tendancielle entre le pourcentage de mots italiens produits dans les énoncés dont le contenu ne porte pas sur l'enfant et ce pourcentage lorsque l'enfant est le thème du tour de parole. En s'adressant aux autres adultes, le père produit 11,6% d'italien dans les tours dont le contenu ne porte pas sur l'enfant, mais lorsqu'il parle de son fils aux autres locuteurs adultes, ses taux d'usage de l'italien atteignent 17.8%.

Bien que cette augmentation ne soit qu'une tendance à confirmer, nous pouvons toutefois supposer que l'évocation du fils dans le discours du père infléchit sa production de l'italien de la même manière que lorsqu'il s'adresse directement à son fils.

3. Discussion

Nos analyses mettent en lumière plusieurs aspects des usages plurilingues au sein du noyau familial de cet enfant grandissant en Vénétie. À partir d'une description globale de la composition codique des tours de parole des adultes, nous avons vu qu'en général, le dialecte est la langue la plus employée, alors que la production de l'italien est plus limitée. En outre, les usages du dialecte et de l'italien chez les parents de l'enfant se situent toujours aux extrêmes de l'échelle : leurs scores d'italien sont les plus élevés et ceux du dialecte sont les plus faibles. Ces premières observations directes des pratiques effectives corroborent les résultats issus des questionnaires de l'*ISTAT* (cf. introduction) qui montrent la vitalité du dialecte dans les échanges quotidiens en situation familiale en Vénétie. Toutefois, elles attestent également de la diffusion croissante de l'italien au sein même de ce type d'interaction.

Dans l'analyse 2, nous avons vu que les adultes réorganisaient la configuration langagière habituelle de leurs tours de parole lorsqu'ils s'adressaient à l'enfant. Tous les adultes augmentent leur production de l'italien au détriment du dialecte. Toutefois cette tendance donne lieu à une variation interindividuelle, puisqu'elle est plus marquée chez les jeunes adultes, dont le père et la mère de l'enfant. Ainsi, malgré la vitalité du dialecte, les parents plurilingues semblent privilégier l'italien pour parler à leur enfant. On peut s'interroger sur la motivation éducative de cette pratique langagière qui semble caractériser les nouvelles générations de parents.

Une autre étude menée dans la même ville (Ghimenton & Chevrot, 2006) apporte des éléments qui permettent d'approfondir notre réflexion. Des interactions en famille ont été enregistrées entre deux fillettes (âgées de 5 et 9 ans respectivement) et les membres de leur entourage familial. La même tendance à favoriser l'italien se manifeste dans les échanges en direction des deux fillettes. Cependant, deux faits nouveaux émergent de cette étude : d'une part, les adultes s'adressent davantage en dialecte à l'aînée ; d'autre part, celle-ci montre des signes de progrès dans l'acquisition du dialecte (la longueur moyenne de ses énoncés dialectaux se situe entre celle de la cadette et celle de sa mère). Ainsi, le comportement éducatif des parents ne viserait pas à éliminer le dialecte du répertoire enfantin, mais à hiérarchiser l'ordre d'acquisition des langues en présence. Compte tenu de la plus grande extension de la langue italienne par rapport au dialecte – aussi bien du point de vue diamésique et diatopique que du point de vue de sa stratification sociale – on peut émettre l'hypothèse que les parents donnent la priorité à l'apprentissage de l'italien afin de faciliter l'insertion des enfants dans des réseaux sociaux élargis, notamment dans le milieu scolaire où la langue utilisée est l'italien. Notre étude a d'ailleurs montré que cette préoccupation se manifesterait particulièrement chez le père, sa vigilance semblant le conduire à activer l'italien même dans des échanges où il évoque son fils sans lui adresser directement la parole.

Les travaux présentés dans cette contribution sont une partie d'une étude visant à mieux comprendre l'influence qu'exerce la société plurilingue sur la construction des compétences langagières. En effet, dans une telle société, la variation engendrée par la généralisation de l'alternance codique augmente la palette des choix communicationnels et stylistiques que l'enfant peut opérer lors de la construction de son répertoire. Cette étude a montré que les adultes modifiaient leur choix de langue quand ils s'adressaient à l'enfant, vraisemblablement en fonction de motivations éducatives. Un suivi longitudinal des

productions de l'enfant est en cours pour observer l'effet de ces stratégies familiales sur la construction de son répertoire plurilingue.

Références

- BERRUTO, Gaetano (1995). *Fondamenti di sociolinguistica*. Roma-Bari : Laterza.
- BERRUTO, Gaetano (2006). Dialect/ standard convergence, mixing, and models of language contact : the case of Italy, in *Dialect change: convergence and divergence in european languages*, Auer, Peter, Hinskens, Frans & Kerswill, Paul (Eds). Cambridge : Cambridge University Press, 81-95.
- DE HOUWER, Annick (1990). *The acquisition of two languages from birth : a case study*. Cambridge: Cambridge University Press.
- DE MAURO, Tullio (1963). *Storia linguistica dell'Italia unita*. Bari: Laterza.
- GHIMENTON, Anna & CHEVROT, Jean-Pierre (2006). Language acquisition in a multilingual society : the case of Veneto, Italy, in *Language variation – European perspective : selected papers from the Third International Conference on language variation in Europe (ICLaVE 3)*, Amsterdam, June 25, Hinskens, Frans. Amsterdam/ Philadelphia : John Benjamins, 71-81.
- GIACALONE-RAMAT, Anna (1995). Code-switching in the context of dialect/ standard language relations, in *One speaker, two languages*. Milroy, Lesley & Muysken, Pieter (Eds). Cambridge, New York, Melbourne : Cambridge University Press, 45-67.
- MARCATO, Gianna (1981). Dialecto e società nel Veneto, in *Guida ai dialetti veneti 111*, Cortelazzo, Manlio (Ed). Padova : Cleup, 61-90.
- SABATINI, Francesco (1997). *L'italiano : dalla letteratura alla nazione. Linee di storia linguistica d'Italia*. Firenze : Accademia della Crusca.
- SANGA, Glauco, (1981). Les dynamiques linguistiques de la société italienne (1861-1980), *Langages*, 61 : 93-115.
- TELMON, Tullio (1990). *Guida allo studio degli italiani regionali*. Alessandria : Edizioni dell'Orso.

LES DÉTERMINANTS DANS L'INTERLANGUE DES BILINGUES TURC-FRANÇAIS ET TURC-ANGLAIS

Jeanne GONAC'H

jeannegonach@yahoo.fr

Dyalang CNRS FRE 2787

8 rue Thomas Becket

76821 Mont-Saint-Aignan Cedex

Abstract : This paper analyses elements of an 'unstable zone' in texts written by Turkish-English and Turkish-French high school and university students. In the first section of this paper, the notion of interlanguage is presented from a historical perspective. The next section describes the methodology of the survey. Finally, initial results are presented of the analysis of determiners as the 'unstable zone' in the corpus.

La littérature sur l'interlangue des apprenants met en évidence le caractère particulièrement instable de ce système. Les chercheurs en linguistique se sont largement intéressés à l'interlangue des apprenants et les applications didactiques de leurs travaux ont déjà bénéficié aux acteurs qui sont au contact d'apprenants. Il n'en reste pas moins un vaste champ à investir tant la question *du contact des langues* et en l'occurrence de l'apprentissage d'une langue seconde ou étrangère est constamment remise à jour par le renouvellement constant des phénomènes migratoires des populations. La multiplication des flux migratoires a en effet entraîné une diversification des populations en contact et donc des langues en contact. De ce perpétuel renouvellement du contact des langues émergent de nouvelles questions sur l'acquisition des langues par des non-natifs. Dans cette perspective, cet article propose de décrire et analyser des éléments de variation dans l'interlangue d'adolescents et d'adultes de minorités sociolinguistiques, (définis par Vermès, 2003) d'origine turque scolarisés en France et en Angleterre.

1. L'instabilité de l'interlangue

Il convient de décrire ici le socle théorique sur lequel a émergé la notion d'interlangue afin de mieux cerner l'orientation des recherches actuelles sur l'interlangue et *a fortiori* notre propre regard sur la question.

1.1. La naissance d'un courant de recherche

Les recherches sur l'interlangue sont nées de l'analyse des erreurs (AE), elle-même issue de l'analyse contrastive (AC). Le principal but de l'AC est d'analyser les divergences et convergences entre les systèmes linguistiques¹ (Tirvassen, 2003). L'AC considère ainsi que l'étude comparée des systèmes phonologiques, syntaxiques, morphologiques etc. de deux langues permet de prévoir les erreurs potentielles d'un individu natif d'une langue 1 apprenant une langue 2. La faiblesse du postulat de ce domaine a rapidement été soulignée puisque peu d'erreurs supposées apparaissaient effectivement dans les productions. En s'opposant au postulat fondateur de l'AC, l'AE s'est intéressée aux erreurs produites. '*The significance of learners' error*' par P.S. Corder en 1967 est l'étude pionnière de ce qui se constituera par la suite comme un domaine de recherche à part entière. La visée de l'AE est

¹ « Contrastive analysis [...] is concerned with the comparison of two or more languages in order to determine both the differences and similarities that hold between them » (Tirvassen, 2003 : 127).

de comprendre et corriger les erreurs (James, 1998 : 235). Selon les chercheurs en AE, les non-natifs font des erreurs qu'un natif ne peut faire. Autrement dit, il existe un phénomène de transfert de la langue d'origine sur la langue du pays d'accueil (LPA) et celui-ci explique que des erreurs soient propres aux non-natifs. A ce jour, l'idée que le phénomène de transfert² est responsable des erreurs dans les productions des non-natifs ne fait pas l'unanimité. Certains chercheurs expliquent les erreurs dans les productions des non-natifs en fonction de la structure de la langue d'origine (Lefrançois, 2001). D'autres montrent qu'il est très délicat de retracer les motifs exacts de l'erreur tant ces motifs peuvent être variés (Tarone, 1988, 2005). James (1998) émet une critique à l'égard de certains chercheurs qui selon lui, interprètent souvent de manière trop subjective leurs résultats. Il explique en effet que certains chercheurs ne prennent pas en compte le fait qu'ils font partie de la communauté linguistique observée et que cette appartenance interfère sur l'interprétation de leur résultat. Il faut retenir de cette critique que l'explication des erreurs varie aussi en fonction du regard du chercheur et de son objectif de recherche.

La question de l'interprétation de l'erreur est une des questions au fondement de notre problématique. Nous tenons d'ailleurs à rappeler ici que nous ne sommes pas native de la langue du pays d'origine (LPO) des enquêtés, le turc, ni de la LPA des turc-anglais, l'anglais, et que cela devra être pris en compte lors de l'interprétation des erreurs.

1.2. La langue de l'apprenant

A partir des différentes critiques adressées à l'AC et à l'AE, certains chercheurs, et notamment Corder, se sont tournés plus spécifiquement vers une nouvelle notion : *l'interlangue*. Dans ce cadre-là, les chercheurs considèrent qu'il ne faut plus raisonner en terme d'*erreur* mais plutôt en terme de *zone d'instabilité*. Étudier l'interlangue c'est étudier la langue de l'apprenant en tant que *système* et non pas en tant que *langue avec des erreurs* par rapport à la langue cible. Le constat que le processus d'acquisition d'une langue génère des 'erreurs' pour n'importe quel apprenant quelle que soit sa langue a incité les chercheurs à développer davantage l'idée d'une « langue de l'apprenant », l'interlangue. Selinker (1991) propose le terme d'interlangue en 1972 pour traiter à la fois les études sur l'acquisition d'une seconde langue et celles sur le bilinguisme qui finalement posent les mêmes questions (Porquier *et al.*, 2005). Les différentes dénominations de l'interlangue témoignent aussi de la vitalité de la question aujourd'hui, comme l'indique Vogel :

« Le terme d'interlangue, proposé par Selinker en 1972, a suscité de nombreuses reformulations : systèmes approximatifs, systèmes intermédiaires, dialectes idiosyncrasiques, grammaire intériorisée ou encore langue de l'apprenant » (cité par Marquillo Larruy, 2003 :73).

Derrière la variété des dénominations, on s'aperçoit que les auteurs ont souhaité insister tour à tour soit sur le fait que l'interlangue tend à se rapprocher du système de la langue cible, soit sur l'idée que l'interlangue est un système évolutif ou encore que l'interlangue est un système propre à l'apprenant.

L'évolution de la réflexion des chercheurs, de l'analyse contrastive à l'interlangue, va de pair avec le développement de la problématique plus globale du contact des langues. Au début des recherches dans ce domaine, le locuteur n'était pas pris en compte dans la réflexion sur le contact des langues. Les chercheurs ne se penchaient alors que sur la question du contact des systèmes linguistiques (Martinez *et al.*, 2000 : 11). Puis, on s'est aperçu que le déni du locuteur, en tant qu'individu psychosocial, empêchait la théorie d'être opératoire. A partir

² Porquier *et al.* (2005 : 25) définissent le transfert comme « un processus psycholinguistique consistant à utiliser dans la construction de l'interlangue des ressources tirées de L1 ».

des années 70, il était désormais impensable de laisser de côté l'apprenant. C'est à ce moment-là que la majorité des chercheurs ont décrit l'apprenant comme un individu pouvant agir sur sa propre langue, et un individu faisant partie d'un environnement spécifique.

1.3. L'instabilité de l'interlangue

L'AC, l'AE, et les recherches sur l'interlangue partagent un même objectif : expliquer les motifs de l'instabilité de la langue de l'apprenant langue seconde³. Dans son dernier ouvrage, Selinker dresse une liste de tous les facteurs susceptibles d'être responsables de l'instabilité de l'interlangue, et parmi ceux-ci : les typologies des langues en contact, les méthodes d'enseignement ou d'apprentissage, les universaux du langage, les stratégies de simplification et complexification. (Selinker, 2005). Ellis (Tarone, 1988 : 19), lui, a choisi de modéliser les différents types de variation de façon binaire : des variations systématiques, d'autres non-systématiques. Par variation systématique, Ellis entend que celle-ci peut être expliquée. Ce type de variation peut-être individuel et dans ce cas il correspond à la motivation et à l'attitude de l'apprenant. L'exemple suivant permet, nous semble-t-il, d'illustrer la variation individuelle proposée par ce modèle : ce sont les propos d'Ihami, un jeune étudiant d'origine turque, rapportés par Becker *et al.* (1988) :

« Lorsqu'il parle allemand, le plus important pour Ihami c'est de maîtriser la situation 'parfaitement'. Plutôt que de réfléchir au choix de l'expression la plus appropriée à la situation donnée, il préfère ne pas choisir et en proposer plusieurs. Il est assez content de son niveau d'allemand et bien qu'il soit conscient de ses problèmes, il n'en est pas embarrassé⁴ » (Becker *et al.*, 1988 : 205 notre traduction)

Les moyens mis en œuvre par Ihami pour interagir 'efficacement' ne sont pas propres à l'ensemble des apprenants mais sont systématiques dans ses productions. Ellis oppose aux variations individuelles, les variations de type contextuelles, et dans ce cas elles peuvent être soit linguistiques soit situationnelles. Enfin, par variation non systématique, Ellis entend toutes les variations que l'on ne peut pas prévoir, que l'on ne peut expliquer ni en corrélation avec d'autres variations ni en fonction de facteurs extra-linguistiques.

2. Le terrain et la problématique

2.1. L'enquête

L'enquête⁵ s'est déroulée auprès de lycéens et d'étudiants monolingues anglais et français et bilingues turc-français et turc-anglais. Nous avons choisi nos interviewés de la manière suivante : les « monolingues » sont nés en France et en Angleterre de parents français et anglais, les bilingues sont scolarisés depuis plus de dix ans en France ou en Angleterre, leurs parents sont turcophones. Nous avons nommé les groupes selon leur origine : les lycéens d'origine turque en France LOTF, les lycéens d'origine française LOF, les étudiants d'origine turque en France EOTF et les étudiants d'origine française EOF ; les lycéens d'origine anglaise LOA, les lycéens d'origine turque en Angleterre LOTA, les lycéens d'origine anglaise EOA et les étudiants d'origine turque en Angleterre EOTA. Pour chaque

³ Nous avons déjà spécifié que nous ne faisons, au même titre que Porquier *et al.*, aucune distinction entre apprenant langue seconde et bilingue.

⁴ Instead of thinking twice which from a certain range of expressions might be the most appropriate, he prefers to offer them all. He is quite satisfied with his knowledge of German ; although he is aware of deficiencies, he is not bothered by them"

⁵ L'enquête a été réalisée dans le cadre du projet ACI dirigé par Akinci, 2003, « Étude psychosociolinguistique du développement de la littéracie chez les bilingues et les monolingues ».

groupe nous avons interviewé environ vingt personnes soit en tout 163 personnes dont 80 bilingues et 83 monolingues.

Les lycéens ont été rencontrés dans leurs lycées, les étudiants se sont soit déplacés dans notre laboratoire pour participer à l'enquête soit nous ont rencontrée dans un lieu de leur choix. L'ensemble du protocole comprenait quatre tâches. Avant de les réaliser, les interviewés ont regardé un film de trois minutes tourné dans une école. Ce film présente des problèmes d'incivilité qui ont lieu quotidiennement : des élèves qui copient les uns sur les autres, des bagarres entre garçons, des filles qui excluent de leur groupe une autre fille, etc. Après avoir visionné ce film, les interviewés devaient remplir un questionnaire. En plus des renseignements sociologiques (sexe, âge, professions des parents, lieu de naissance), des questions portaient sur les pratiques linguistiques (pratique de la LPA et de la LPO avec la famille et les amis) et sur les pratiques littéraciques (lecture, écriture dans la LPA et la LPO) des enquêtés. Puis, les enquêtés devaient produire quatre textes : un texte narratif écrit, un texte narratif oral, un texte expositif écrit et un texte expositif oral.

2.2. Questionnement et visée

L'objectif de l'enquête est double. D'une part, il s'agit de reconnaître s'il existe des convergences dans l'interlangue des bilingues turc-anglais et turc-français. D'autre part, il est question de comprendre si le croisement des résultats du questionnaire avec des éléments de la production scripturale et orale des enquêtés révèle des tendances significatives. L'hypothèse est que les bilingues qui déclarent lire et écrire souvent dans la LPA sont aussi ceux dont les textes comportent le moins d'écart⁶.

L'intérêt de notre étude est double. Tout d'abord, alors que les études qui portent sur l'interlangue des apprenants s'intéressent très souvent à la période des premiers contacts avec la LPA (Lefrançois, 2001), nous nous sommes penchée sur des apprenants qui sont scolarisés depuis plus de dix ans dans le pays d'accueil. Nous souhaitons ainsi observer s'il reste des traces de l'instabilité de l'interlangue après plus de dix de scolarisation dans le pays d'accueil. D'autre part, comme nous nous intéressons à des bilingues qui partagent la même langue d'origine mais dont la LPA diffère, nous cherchons à donner des éléments de réponse quant à l'influence de la langue d'origine sur les productions dans la LPA.

3. Une zone instable dans les textes des turcs-français et des turcs-anglais : les 'petits mots fonctionnels'

L'analyse des textes de notre corpus a révélé une zone d'instabilité commune aux quatre groupes de bilingues (LOTF, EOTF ; LOTA, EOTA). Par conséquent, nous analyserons uniquement les textes des bilingues dans cet article. Cette zone d'instabilité est constituée de 'petits mots fonctionnels' et particulièrement de déterminants, de pronoms et de prépositions. La difficulté des apprenants par rapport à l'emploi de ces mots a été soulignée dans d'autres études qu'il convient de présenter.

3.1. Les 'petits mots' marqueurs de l'instabilité

Dans une enquête menée par Krashen (Tarone, 2005), une des enquêtées, une femme d'origine chinoise apprenant l'anglais, explique qu'elle a décidé de ne pas porter attention aux 'petits mots fonctionnels', en l'occurrence, les prépositions *to*, et *of*, les articles définis et indéfinis *the* et *a* (Tarone, 2005 : 160). Pour l'enquêtée, l'utilisation de ces 'petits mots' n'est

⁶ Comme nous nous positionnons au sein des réflexions récentes sur l'interlangue, il apparaît difficile d'utiliser le terme d'« erreur » pour désigner les marqueurs de l'instabilité de l'interlangue (cette dernière expression étant trop longue). Nous avons choisi « écart » d'après Delamotte-LeGrand (2005).

pas indispensable pour se faire comprendre. Dans une autre étude, à propos de productions d'apprenants de français langue étrangère (FLE), Chiss *et al.* (1984) soulignent :

« Il apparaît qu'en matière d'écrits en FLE produits par des apprenants de niveau universitaire, l'inventaire des difficultés inclut encore le plus souvent l'ensemble des composantes linguistiques : la non-maîtrise des phénomènes grammaticaux intra-phrastiques (emploi des prépositions, des déterminants, place des adjectifs épithètes etc. » (Chiss et al., 1984 : 28).

Nous pouvons avancer que les prépositions, les pronoms, les déterminants sont des marqueurs de l'instabilité de l'interlangue dans la mesure où ils ont déjà été repérés à plusieurs reprises dans différents corpus de productions par des non natifs. Les écarts liés à l'utilisation de ces mots sont soit des ellipses soit des substitutions.

Il convient à présent d'aborder la question des motifs de ces écarts. Une première explication qui s'offre à nous est l'idée que cette zone d'instabilité est liée aux caractéristiques phonologiques de ces mots. *De, à, of* etc. sont des mots monosyllabiques, généralement non accentués ; à l'oral, la substitution ou l'ellipse d'un de ces mots peut en effet ne pas poser de problème d'intercompréhension. Une autre thèse avancée par Young (1996) montre que les écarts liés à ces mots sont dus à leur sens.

« La difficulté à apprendre les relations complexes entre forme et fonction a été soulignée par Anderson (1984), qui a proposé dans son principe un-pour-un que l'acquisition d'une nouvelle forme est facilitée quand il existe un rapport clair et unique entre la forme et le sens' » (Young, 1996: 142, notre traduction).

Or en français et en anglais, le rapport entre la forme des articles et leur sens n'est ni clair ni unique mais complexe. Dans la suite de son article, Young affirme que les articles en anglais constituent une zone d'instabilité pour des apprenants de langue étrangère dans le cas où les déterminants ne sont pas des formes attestées dans les langues de ces apprenants. L'influence de la langue d'origine serait donc l'une des raisons pour laquelle les apprenants ne maîtrisent pas ces mots. En turc, les déterminants n'existent pas. Doit-on pour autant conclure que la zone d'instabilité attestée dans les textes des bilingues est uniquement liée à cette absence de déterminants dans la langue d'origine ?

3.2. Des 'petits mots' variés dans notre corpus

3.2.1. Inventaire des marqueurs de l'instabilité de notre corpus

Nous avons recensé 150 'petits mots' ou 'phénomènes grammaticaux intraphrastiques' dans les textes de 70 scripteurs bilingues. Le tableau 1 ci-dessous donne une vision globale des écarts liés à ces mots.

⁷ « The difficulty of learning complex form-function relations was stressed by Anderson (1984), who proposed in the One to One Principle that acquisition of a new form is facilitated when there is a clear and unique correspondence between form and meaning » (Young, 1996: 142).

Nombre d'écarts liés à l'usage des 'petits mots fonctionnels'				
	Les déterminants	Les prépositions	Les pronoms	Total
LOTF	9	13	11	33
EOTF	10	11	10	31
LOTA	17	12	13	42
EOTA	20	13	7	40
Total	56	49	41	146

Tableau 1 - Nombre d'écarts liés à l'usage des 'petits mots fonctionnels'

Le tableau indique que les écarts par rapport à l'utilisation de ces mots sont plus nombreux pour les deux groupes bilingues turc-anglais que les deux groupes de bilingues turc-français. Le cas des déterminants illustre particulièrement ce constat. C'est cette catégorie de petits mots que nous avons choisi de décrire dans cet article. Nous en avons recensé 56 dans les textes de 35 scripteurs différents.

3.2.2. Des difficultés à catégoriser les déterminants

Voici le texte d'un lycéen turc-anglais dans lequel nous avons relevé un écart au niveau du choix du déterminant.

- 1. One day I was in wood green with my friends and there was some thomas more boys and girls they they started to bully my friend. So me and my other friends went to help him, then we had a big fight against them, then some girl came and punched my friends girl friend then my friends got **the** glass bottle and broke it on her head.

Le déterminant 'the' n'a pas été introduit dans l'univers du discours auparavant et aucun élément du texte ne nous permet de référer à cette 'bouteille' en particulier. C'est l'article indéfini, 'a' qui était attendu ici. La difficulté à choisir l'un des deux articles a déjà été soulignée dans l'article de Young (1996) :

« L'article défini, the et l'indéfini, a, fonctionnent en anglais de façons diverses qui se chevauchent entre elles ⁸ » (Young, 1996 :135).

C'est le système de référence qui pose problème, mais pas seulement. Le choix de ce déterminant peut s'expliquer de différentes manières. On peut en effet supposer que le scripteur emploie ce déterminant en pensant que l'article défini ici va lui permettre de référer à un élément spécifique. On peut aussi penser que le scripteur visualise la scène et fait référence au contexte situationnel et ainsi emploie 'the' comme un déictique. Comme il a certainement déjà raconté cette histoire, il s'attend à ce que l'interlocuteur sache que le blessé a été frappé par une bouteille en verre.

- 2. One day on the road near [] supermarket

Dans cet exemple, le lycéen ne détermine pas le nom *supermarket* alors qu'il aurait fallu spécifier de quel 'supermarché' il s'agissait. Comme nous ne pouvons pas revenir avec le scripteur sur cette erreur, et qu'aucun autre exemple dans les textes ne nous permet de trancher sur ce point, nous proposons plusieurs interprétations :

- L'ellipse du déterminant est liée au fait que, pour le scripteur, la préposition *near* permet de déterminer le nom suivant. Selon cet élève, le rôle joué par la préposition suffit pour déterminer le nom *supermarket*, comme c'est le cas dans l'exemple suivant : On the road near London. Dans ce cas, on peut penser que c'est une surgénéralisation de la règle qui précise qu'on ne détermine pas les noms de lieu.

⁸ « The definite article, the and the singular, a(n), in English function in a variety of ways that overlap with each other linguistic forms » (Young, 1996 :135).

- Le déterminant qui introduit 'road' suffit à introduire tout le syntagme nominal « **The** road near supermarket ».

Nous avons voulu souligner à travers l'analyse de ces deux exemples la multiplicité des motifs possibles pour expliquer l'ellipse ou la substitution d'un déterminant dans le corpus. Il s'avère très difficile de classer les 54 autres exemples car les écarts liés à ces mots peuvent porter à la fois sur les règles du genre, du nombre, du type de référence, de la distinction générique / spécifique, du type d'élément déterminé (animé / non animé) etc.

3.2.3. Des corrélations muettes

Le fait de ne pas réussir à dégager des règles sous-jacentes à ces écarts nous a conduit à les analyser sous un autre angle. Les recherches de Labov (1976, 1978) ont montré des corrélations entre certaines variations et le milieu social des enquêtés. Nous avons voulu tester l'hypothèse des corrélations entre les variables choisies dans notre questionnaire et les variations au niveau des déterminants. Notre questionnement était le suivant : est-ce que les écarts trouvés au niveau de l'emploi des déterminants sont liés au fait que les enquêtés déclarent lire plus fréquemment dans la LPA ou la LPO que d'autres ?

Les résultats de nos analyses ne permettent pas non plus de faire émerger des tendances entre les écarts et les éléments relevés dans le questionnaire. S'ajoute donc à la difficulté de classer ces écarts le manque de corrélations que nous pouvons établir entre ces derniers et les pratiques linguistiques et littéraires de chaque enquêté. Nos résultats font écho à une réflexion de Tabouret-Keller (1985) :

« De l'extrême représenté par des corrélations qui s'imposent comme une sorte d'impératif social à celui représenté par la quasi-imprévisibilité du locuteur, il y a un champ immense dont il ne faut pas s'étonner qu'il échappe à la visée de la sociolinguistique » (Tabouret-Keller, 1985 : 128).

Si les corrélations sont 'muettes' entre les variations scripturales des enquêtés et les variables testées c'est peut-être aussi que les variables choisies ne suffisent pas à donner des résultats sur ces croisements. Sur ce point, il nous semble en effet que le fait que l'enquêtrice ne soit pas native de l'anglais a eu une influence sur la façon dont les enquêtés bilingues turc-anglais se sont comportés. Nous rejoignons ici les hypothèses des recherches en interaction (Vasseur, 2000, Tarone, 1988). La *pression normative*, moins élevée dans le cas des enquêtes en Angleterre, a pu affecter la façon dont les enquêtés bilingues ont prêté attention à leur expression. Cette différence entre le corpus en France et en Angleterre est, nous semble-t-il, un des facteurs responsables du plus grand nombre d'écarts dans le corpus des bilingues turc-anglais que dans celui des bilingues turc-français.

3.2.4. Des pistes pour un classement

On s'est aperçu que les écarts et particulièrement les ellipses sont relevés lorsque deux petits mots se suivent sur l'axe syntagmatique. Par exemple, après une préposition, nous avons recensé treize variations sur des déterminants, dont l'exemple suivant :

3. I think they was fighting **for** () girl

On retrouve cette difficulté à combiner les petits mots lorsqu'ils se trouvent après une conjonction de coordination. Ces exemples constituent une piste à explorer à travers d'autres corpus.

4. It was so painful and [] most painful issue was when other students laughed at me.
5. It's wrong and [] bad thing
6. Ils ont des lunettes ou [] appareil dentaire.
7. ...Et plus particulièrement le non ou [] manque de respect à l'égard du professeur.

4. Conclusion

Notre étude révèle qu'il existe une zone d'instabilité dans l'interlangue des minorités sociolinguistiques scolarisées depuis plus de dix ans dans le pays d'accueil. Des convergences dans les écarts au niveau des déterminants ont effectivement été mises au jour dans les productions des quatre groupes de bilingues de notre corpus. Cependant, la diversité de ces écarts ne nous permet pas d'affirmer que la langue d'origine a une influence sur ces derniers. Nous avons montré que de multiples facteurs pouvaient être responsables de l'apparition de ces écarts. Afin de progresser sur la question des motifs des écarts dans les productions des apprenants, nous proposons de prendre en compte les deux points suivants. D'une part, il faudrait intégrer en amont de l'enquête de nouvelles variables (identité de l'enquêté, entre autres). D'autre part, il faudrait multiplier les exemples et tester les variables sur de nouveaux corpus.

Références bibliographiques

- BECKER, Angelica, CARROL, Mary, KELLY, Ann (Eds) (1988). *Reference to space*. Strasbourg & Heidelberg.
- CHISS, Jean-Louis, FILLIOLET, Jacques (1984). Production écrite et apprentissage des types de discours en français langue étrangère. *Linx*, 11 : 27-48.
- DELAMOTTE-LEGRAND, Régine (2005). Mal dit – mieux dit – bien dit – inédit : un point de vue sociolinguistique sur l'évaluation, *Repères*, 31 : 141-160.
- JAMES, Carl (1998). *Errors in language learning and use*. London : Longman.
- LABOV, William (1976). *Sociolinguistique*. Paris : Les éditions de Minuit.
- LABOV, William (1978). *Le parler ordinaire*. Paris : Les éditions de Minuit.
- LEFRANÇOIS, Pascale (2001). Le point sur les transferts dans l'écriture en langue seconde, *Canadian Modern Language Review*, 58, 2 : 223-245.
- MARQUILLO LARRUY, Martine (2003). *L'interprétation de l'erreur*. Paris : Clé International.
- PORQUIER, Rémy, PY, Bernard (2005). *Apprentissage d'une langue étrangère : contexte et discours*. Paris : Didier.
- MARTINEZ, Pierre, PEKAREK DOELHER, Simona (2000). *La notion de contact des langues en didactique*. Fontenay Saint-Cloud : ENS éditions.
- SELINKER, Larry (2005). Fossilization or 'does your mind mind ?', in *Studies of fossilization in second language acquisition*. Han, Zhao Hong & Odlin, Terence (Eds). Clevedon: Multilingual matters, 201-210.
- SELINKER, Larry (1991). *Rediscovering interlanguage*. London/ New York : Longman.
- TABOURET-KELLER, André (1985). Langage et société : les corrélations sont muettes. *La linguistique*, 21: 125-139.
- TARONE, Elaine (1988). *Variation in Interlanguage*. London : Edward Arnold.
- TARONE, Elaine (2005). Fossilization, social language and language play, in *Studies of fossilization in second language acquisition*, Han, Zhao Hong & Odlin, Terence (Eds). Clevedon : Multilingual matters, 157-172.
- TIRVASSEN, Rada (2003). Approcher les contacts des langues à partir des outils issus du structuralisme, quel usage effectuer du terme interférence ?. *Glottopol*, 2 : 123-133.
- VASSEUR, Marie-Thérèse (2000). *L'interaction interlingue, dimension dialogique, reflexive et temporelle*. Habilitation à diriger des recherches, vol. 1, Paris : Université René Descartes.
- VERMÈS, Geneviève (1990). L'entrée dans l'écrit des minorités sociolinguistiques. *Migrants-Formation*, 83 : 65-74.
- YOUNG, Richard (1996). Form-function relations in articles in English Interlanguage, in *Second language acquisition and linguistic variation*, Bayley, Robert & Preston, Dennis. (Eds). Amsterdam : John Benjamins, 135-175.

USAGES LANGAGIERS ENFANTINS ET VARIABLES SOCIOLINGUISTIQUES : CONVERGENCES MICROSOCIOLOGIQUES ET STRATIFICATION MACROSOCIOLOGIQUE¹

Aurélie NARDY, Nathaël MARTIN

Aurelie.Nardy@u-grenoble3.fr, Nathael.Martin@u-grenoble3.fr

Laboratoire Lidilem – Université Stendhal

1180, Avenue Centrale – BP 25

38040 Grenoble cedex 9

Abstract : This article presents three studies (1 macro-sociological and 2 micro-sociological) which explore the use of sociolinguistic variables among young children. The results suggest a scenario which takes into account linguistic convergences at a micro-sociological level and social stratification of the uses at a macro-sociological level.

Introduction

Dans le domaine de la sociolinguistique, de nombreuses enquêtes à grande échelle ont montré que les usages langagiers des locuteurs étaient socialement stratifiés en fonction de différents facteurs extralinguistiques comme le milieu social, l'âge, le sexe, etc. Cette tendance macrosociologique a été établie chez les locuteurs adultes dès les premières études sociolinguistiques, notamment par Labov (1976), et plus récemment chez les enfants (Roberts, 1994, 1997; Chevrot, Beaud & Varga, 2000). D'autres études, à un niveau plus microsociologique, se sont intéressées au lien entre pratiques linguistiques et relations sociales entre les membres d'une même communauté. Ainsi, Milroy (1987) a mis en évidence qu'un réseau social dense et multiplexe², indiquant une forte intégration sociale dans la communauté locale, favorisait l'usage des variantes vernaculaires du groupe. Autrement dit, au niveau microsociologique, plus un locuteur a de liens sociaux avec les membres de sa communauté, plus il utilisera les variantes de son groupe. Russel (1982) a d'ailleurs montré que la conversation entre amis entraînait des phénomènes de convergence dans les usages linguistiques. Ainsi, alors que les études macrosociologiques font ressortir des divergences dans les usages, les études microsociologiques soulignent leur convergence.

A notre connaissance, aucune étude empirique n'a cherché à réunir des résultats qui concilieraient ces deux points de vue. Nous aborderons cette question à travers une démarche développementale. A partir d'une approche macrosociologique, nous verrons à quel âge les usages enfantins divergent. Plus précisément, nous observerons à partir de quel âge se mettent en place les régularités sociolinguistiques observées chez les adultes. Parallèlement, l'approche microsociologique de deux groupes de jeunes enfants nous permettra d'observer si des phénomènes de convergence langagière émergent au bout d'une année de fréquentation quotidienne. En nous appuyant sur les résultats issus de ces différentes

¹ Cette recherche est soutenue par l'ACI Systèmes Complexes en SHS (*Dynamique des interactions sociales et structuration des usages linguistiques au sein des réseaux sociaux : de l'individu au groupe de pairs*) et par l'ACI Espaces et Territoires (*La dynamique spatiale des usages langagiers : territoires, identités et développement linguistique en situation de plurilinguisme*).

² Un individu appartient à un réseau dense si les membres de son réseau se fréquentent aussi entre eux. Son réseau est multiplexe lorsqu'il fréquente les mêmes individus dans ses différents cadres de vie (travail, loisirs, famille, etc.).

enquêtes, nous chercherons à expliquer comment les convergences locales et les divergences globales se combinent au cours du développement. Plus particulièrement, les phénomènes de convergence et divergence seront examinés à travers l’usage enfantin des liaisons facultatives et de son évolution avec l’âge. En effet, ce type de liaison est connu pour être un marqueur sociolinguistiquement différentiel particulièrement bien décrit chez l’adulte.

1. La liaison

La liaison consiste en l’apparition d’une consonne entre deux mots (mot1 et mot2) qui s’enchaînent à l’oral ; par exemple apparition de /n/ entre *un* et *avion* dans [œnavjõ]. Pour que la consonne de liaison apparaisse, le mot2 doit commencer par un son vocalique lorsqu’il est produit en isolation. En revanche, cette consonne n’est jamais présente en finale du mot1 lorsqu’il est situé à la fin d’un énoncé, ni à l’initiale du mot2 lorsqu’il occupe la première position d’un énoncé (ex : *j’en mange un* [ʒãmãʒœ], *avion* [avjõ]). De plus, la nature de la consonne de liaison est déterminée par le mot1, par exemple *un* [œ] active la consonne de liaison /n/, *gros* [gro] active /z/ et *petit* [pɛti] active /t/.

Deux catégories de liaisons existent : les liaisons obligatoires et les liaisons facultatives. Dans notre étude, nous nous appuyons sur le classement proposé par Booij et De Jong (1987). A partir d’un corpus de 38 locuteurs adultes répartis en 3 classes d’âges, 5 classes socio-économiques et avec une répartition équivalente d’hommes et de femmes, ils ont établi les contextes dans lesquels les liaisons étaient réalisées de manière invariable quels que soient le locuteur et/ou ses caractéristiques sociologiques. Ainsi, 4 contextes seulement apparaissent comme relevant des liaisons obligatoires (100% de réalisation). Quant aux contextes de liaison ne conduisant pas à une réalisation systématique, ils ont été classés dans la catégorie des liaisons facultatives ; les auteurs en dénombrent 6 (voir Tableau 1).

Liaisons obligatoires	Liaisons facultatives
- déterminant + X ³ (<i>un ours</i> [œnurs])	- nom pluriel + X (<i>des rideaux anciens</i> [derido(z) ⁴ äsjẽ])
- pronom personnel + verbe (<i>ils arrivent</i> [ilzaRiv])	- adjectif + X (<i>des jeunes amis</i> [deʒœn(z)ami])
- verbe + pronom personnel (<i>vient-il</i> [vjẽtil])	- verbe + X (<i>ils mangent ensemble</i> [ilmãʒ(t)äsäbl])
- expressions figées avec liaison (<i>tout à coup</i> [tutaku])	- auxiliaire + X (<i>ils sont arrivés</i> [ilsõ(t)arive])
	- invariable monosyllabique + X (<i>très intéressant</i> [trẽ(z)ẽteresã])
	- invariable polysyllabique + X (<i>pendant un</i> [pãdã(t)œ])

Tableau 1 - Classement des contextes de liaisons obligatoires et facultatives (Booij & De Jong, 1987)

Comme nous l’avons évoqué plus haut, les liaisons facultatives sont un lieu de variation sociale. En effet, des travaux descriptifs ont montré que leur emploi par les locuteurs adultes variait en fonction du milieu social (Delattre, 1966 ; Malécot, 1975 ; Ashby, 1981 ; Booij & De Jong, 1987 ; De Jong, 1994), de l’âge (Ashby, 1981), du sexe (Ashby, 1981 ; Green & Hintze, 1990 ; De Jong, 1991) et de la situation de communication (Ågren, 1973 ; Moisset,

³ Les "X" représentent, selon les cas, un nom, un adjectif, un déterminant, etc. commençant par une voyelle.

⁴ La consonne entre parenthèses indique la consonne de liaison. Dans le cas des liaisons facultatives, celle-ci peut être réalisée ou non.

2000). Par exemple, les locuteurs qui exercent une profession nécessitant un niveau d'études élevé réalisent davantage de liaisons facultatives que des locuteurs qui ont un niveau d'études plus bas. Les premiers auront tendance à réaliser la liaison facultative dans la séquence « ils sont arrivés » et diront donc [ilsɔ̃tarive] alors que les seconds auront plus tendance à ne pas faire la liaison et à dire [ilsɔ̃arive].

Nous observerons l'usage enfantin des liaisons facultatives dans deux types d'études : une étude transversale à grande échelle que nous appellerons macrosociologique et deux suivis longitudinaux de petits groupes d'enfants se fréquentant régulièrement que nous qualifierons de microsociologiques.

2. Étude transversale macrosociologique

2.1. Échantillon

L'étude transversale macrosociologique concerne un échantillon de 187 enfants âgés de 2 à 6 ans (93 filles et 94 garçons), tous monolingues français, que nous avons répartis en quatre tranches d'âge (2-3 ans, 3-4 ans, 4-5 ans et 5-6 ans) et en deux milieux sociaux nettement contrastés : un milieu dit « parents cadres » et un milieu dit « parents ouvriers ». Les enfants ont ainsi été choisis *a priori* en fonction de la profession de leurs deux parents, ce qui nous a permis de sélectionner seulement les enfants dont le profil social était typique d'un milieu. Afin d'établir deux groupes sociaux distincts, nous avons utilisé la nomenclature INSEE de 1982 (Desrosières & Thévenot, 1988). Les enfants dont la profession des deux parents, ou la profession d'un seul parent si l'autre était sans emploi, s'apparentait au groupe 3 de la nomenclature INSEE (ex : cadres de la fonction publique, professeurs et professions scientifiques, cadres administratifs et techniques, ingénieurs, etc.) ont été considérés comme faisant partie du groupe « parents cadres ». Pour le groupe « parents ouvriers », nous avons choisi des enfants dont la profession des deux parents, ou la profession d'un seul si l'autre était sans emploi, faisait partie du groupe 6 de la nomenclature INSEE (tous les types d'ouvriers et les chauffeurs). Le tableau 2, ci-dessous, indique la répartition des sujets en fonction de la tranche d'âge et du milieu social.

Tranche d'âge	Milieu social	N
2-3 ans	parents cadres	22
	parents ouvriers	18
3-4 ans	parents cadres	25
	parents ouvriers	20
4-5 ans	parents cadres	27
	parents ouvriers	27
5-6 ans	parents cadres	25
	parents ouvriers	23

Tableau 2 - Étude macrosociologique : répartition des enfants en fonction de la tranche d'âge et du milieu social

2.2. Démarche empirique

Afin de solliciter, auprès d'un grand nombre d'enfants, la production extensive de liaisons, nous avons élaboré une tâche expérimentale de dénomination d'images. Cette tâche suscitait la production de liaisons facultatives en contexte "adjectif + nom", contexte facultatif selon la classification de Booij et De Jong (1987). Les mots1 (adjectifs) étaient *petit* et *gros* et les mots2 (noms à initiale vocalique) étaient *ours*, *arbre*, *avion*, *éléphant*, *escargot* et *ordinateur*. Ainsi, chaque animal/objet était représenté une fois en un exemplaire de petite taille (pour susciter la production de *petit* + nom) et une autre fois en un exemplaire de grande taille (pour susciter la production de *gros* + nom). Enfin, dans le but de pallier d'éventuelles influences entre liaisons produites, chaque présentation d'image-cible était

suivie d'une image-distracteur conduisant à la production de séquences sans liaison de type mot1 (*petit, gros*) + mot2 à initiale consonantique (*ballon, lit, cochon, singe, camion, balai*). Au final, chaque enfant était amené à produire 12 séquences-cibles mot1 + mot2 en contexte de liaison facultative "adjectif + nom".

2.3. Résultats

Dans ce paragraphe, nous présenterons l'évolution des pourcentages de liaisons facultatives réalisées conformément à la cible adulte, qui correspondent à la réalisation de la consonne de liaison adéquate, par exemple quand l'enfant a produit [pøtiturs] ou [grozurs]. C'est donc la capacité normative enfantine que mesurent ces pourcentages.

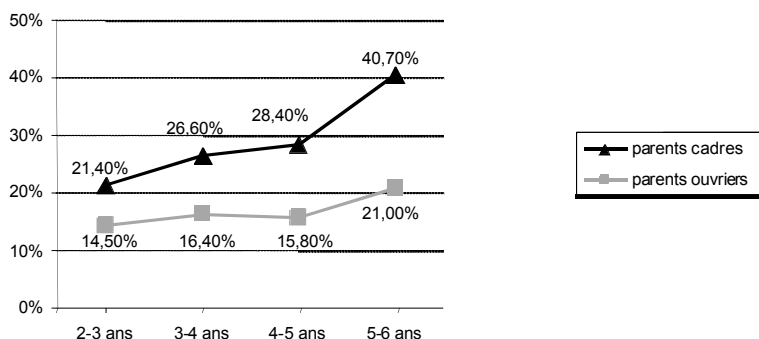


Figure 1 - Pourcentages de liaisons facultatives réalisées conformément à la cible adulte

L'observation de la figure 1 montre que, quelle que soit la tranche d'âge, les enfants issus du milieu cadre réalisent davantage de liaisons facultatives conformément à la cible adulte que les enfants issus du milieu ouvrier. Nous notons d'ailleurs un effet significatif du milieu social ($ANOVA^5$, $F_{1,179} = 11,123$, $p = 0,001$). Nous remarquons également qu'à 2-3 ans, les différences entre les enfants des deux milieux sont peu importantes et qu'elles s'accroissent progressivement durant le développement. A 5-6 ans, la différence entre les taux de réalisation de liaisons facultatives chez les enfants des 2 milieux (40,70% chez les enfants de parents cadres et 21% chez les enfants de parents ouvriers) devient significative ($Holm-Sidak$, $p = 0,011$).

Nous avons également considéré l'évolution de la variance des pourcentages de liaisons facultatives réalisées conformément à la cible adulte. La variance (S^2) est une mesure statistique permettant d'évaluer la dispersion des scores autour de la moyenne⁶. De 2-3 ans à 5-6 ans, la variance augmente (voir tableau 3) ; elle passe de 335,76 à 2-3 ans à 934,93 à 5-6 ans. Le rapport entre la variance à 5-6 ans et la variance à 2-3 ans est d'ailleurs significatif comme l'atteste le calcul du rapport F ($F_{39,47} = 2,78$, $p = 0,00070$) ; nous pouvons alors considérer que ces deux variances sont différentes. Autrement dit, entre 2 et 6 ans, les performances des enfants se distinguent de plus en plus, ce qui conduit à l'apparition des différences entre les enfants des deux milieux (et sans doute à d'autres différences qui ne

⁵ L'ANOVA à 2 facteurs (âge x milieu social) a été effectuée après transformation des données en arc sinus (Sokal et Rohlf, 1981).

⁶ La variance d'une distribution correspond à la moyenne des carrés des écarts par rapport à la moyenne. Plus la variance est grande, plus la dispersion est grande et inversement.

sont pas observées ici⁷). Par ailleurs, l'écart observé entre les deux groupes d'enfants à 5-6 ans laisse présager les différences sociales observées chez les locuteurs adultes.

Tranche d'âge	Pourcentages de liaisons facultatives réalisées	S ²
2-3 ans	18,30%	335,76
3-4 ans	22,10%	491,01
4-5 ans	22,10%	811,79
5-6 ans	31,30%	934,93

Tableau 3 - Évolution des pourcentages de liaisons facultatives réalisées conformément à la cible adulte et de la variance

Au niveau macrosociologique, l'enquête transversale menée auprès d'un grand nombre de sujets met donc en évidence une divergence des usages des enfants des deux milieux entre 2 et 6 ans. La section suivante (section 3) s'attachera quant à elle à rendre compte de l'évolution des usages de deux groupes d'enfants dans une perspective microsociologique.

3. Études longitudinales microsociologiques

3.1. Étude A

3.1.1. Échantillon et démarche empirique

La première étude longitudinale que nous présentons ici concerne le suivi d'un groupe de 14 enfants (5 filles et 9 garçons) de Moyenne Section de Maternelle d'une école de Rennes, âgés en moyenne de 4;6 au premier temps d'observation (T1) à 5;2 au deuxième temps d'observation (T2). Ce groupe présente une composition sociale hétérogène : 8 enfants de parents cadres et 6 enfants de parents ouvriers. Le suivi de cette classe a donné lieu au recueil d'un corpus langagier au sein de la classe pendant la période d'accueil du matin caractérisée par des jeux libres et un contrôle allégé de la maîtresse. Cette période apparaissait propice à l'émergence d'échanges verbaux et de comportements sociaux « spontanés » entre les enfants (Barbu, 2000). Quinze minutes d'enregistrement ont été effectuées pour chaque enfant et pour chaque temps d'observation. L'enfant pour lequel on souhaitait recueillir des données langagières portait un sac à dos équipé d'un micro sans fil afin de le laisser libre de ses mouvements, de minimiser l'effet de l'observation et d'obtenir ainsi des énoncés dans une situation la plus informelle possible. Notons également qu'une période d'habituation d'un mois, avant le recueil de données, a permis aux enfants de se familiariser avec l'observateur et ses outils d'enregistrement.

3.1.2. Résultats

Comme pour l'étude transversale présentée plus haut, nous avons procédé à l'examen de la variance des pourcentages de liaisons facultatives réalisées conformément à la cible adulte⁸.

⁷ Toutefois, d'autres analyses prenant appui sur la variable sexe ont déjà montré que ce facteur n'influençait pas la production enfantine des liaisons facultatives (Nardy, en préparation).

Tps Obs	Enfants														M	S ²	[Min-Max]
	Ale	Ali	Ann	Car	Cor	Fél	Jul	Léa	Mau	Max	Rom	Sop	Tom	Yan			
T1	0	4,76	0	10	0	23,08	0	4,65	0	5,26	0	5	0	0	3,77	40,83	[0-23,08]
T2	0	5,41	3,45	6,52	6,06	2,63	2,56	11,76	2,27	3,57	2,78	0	3,03	0	3,57	9,86	[0-11,76]

Tableau 4⁹ - Étude A : Évolution des pourcentages de liaisons facultatives réalisées conformément à la cible adulte

Comme le montre le tableau 4, nous observons que la variance (S²) diminue significativement de T1 à T2¹⁰ ($F_{13,13} = 4,14$, $p = 0,00778$) tout comme l'écart entre le pourcentage minimum et le maximum ([Min-Max]). Les usages des enfants de la classe semblent donc converger au bout de huit mois de fréquentation quotidienne.

Les études, transversale et longitudinale, montrent des résultats différents. Dans l'étude transversale, au niveau macrosociologique, nous avons vu que les pourcentages de réalisation des liaisons facultatives tendaient à diverger alors que dans cette étude longitudinale, au niveau microsociologique, les pourcentages tendent à converger malgré la diversité sociale du groupe.

Nous présenterons maintenant une deuxième étude longitudinale microsociologique menée auprès d'enfants d'âges comparables.

3.2. Étude B

3.2.1. Échantillon et démarche empirique

Le suivi longitudinal de l'étude B concerne un groupe de 11 enfants (5 filles et 6 garçons), d'origines sociales contrastées (6 enfants de parents cadres et 5 enfants de parents ouvriers) de la même classe de maternelle de la banlieue grenobloise. Les enfants ont été suivis sur une période d'un an, de la fin de la Moyenne Section à la fin de la Grande Section de Maternelle. Deux temps d'observation ont été menés ; la moyenne d'âge des enfants au T1 était de 4;7 et de 5;7 au T2. Ils ont été enregistrés individuellement pendant l'atelier semi-dirigé, moment de la journée de classe durant lequel les enfants pratiquent une activité dans le groupe de pairs. Cet atelier durait entre 30 et 45 minutes, soit autant de temps d'enregistrement et de corpus transcrit pour chacun des enfants. Enfin, ces derniers ont été enregistrés dans les mêmes conditions que dans l'étude A.

3.2.2. Résultats

Dans l'étude B, notre intérêt s'est porté sur l'observation conjointe de plusieurs variables connues pour être sociolinguistiquement différentielles chez l'adulte et l'adolescent : la liaison facultative, le /R/ en position post-consonantique finale (Laks, 1977) et le /l/ dans les clitiques sujets « il(s) » et « elle(s) » (Laks, 1980). Pour ces 3 variables, les locuteurs issus d'un milieu « favorisé » auront davantage tendance à réaliser le segment consonantique. Par

⁸ Notons que les pourcentages de liaisons facultatives réalisées de l'étude A sont inférieurs à ceux relevés dans l'étude précédente (étude macrosociologique). En situation expérimentale, seul le contexte *adjectif + nom* était présent ; de plus il était limité à des liaisons après les mots *petit* et *gros* connus pour favoriser la réalisation de la liaison chez l'adulte. Ici, la situation écologique de recueil des données nous a permis de relever tous les contextes de liaisons facultatives produits par les enfants.

⁹ Les valeurs des tableaux 4 et 5 sont toutes exprimées en pourcentages à l'exception de la colonne S².

¹⁰ Notons toutefois que l'ensemble de la convergence repose uniquement sur *Fél* dont les pourcentages de liaisons facultatives réalisées chutent considérablement de T1 à T2 (23,08% au T1 contre 2,63% au T2). L'étude microsociologique B viendra étayer la convergence observée dans l'étude A.

exemple, ces locuteurs diront plus souvent [dävātø] que [dävāø] (*devant eux*), [etrkōtā] que [etkōtā] (*être content*), [ilfēbo] que [ifēbo] (*il fait beau*).

Les résultats que nous présentons ci-après concernent les pourcentages de variantes standard produites par les enfants pour ces trois variables.

Comme pour l'étude A, nous constatons que la variance (S^2) diminue de manière significative ($F_{10,10} = 3,11$, $p = 0,04390$) entre les deux temps longitudinaux de la même façon que l'écart entre les valeurs minimale et maximale (voir tableau 5). Il apparaît alors qu'au sein du même groupe d'enfants, d'une année scolaire à l'autre, les usages langagiers convergent. Les enfants dont les pourcentages de variantes standard étaient élevés ont tendance à les diminuer du premier au second temps d'observation et vice-versa.

Tps Obs	Enfants											M	S ²	[Min-Max]
	Ale	Ama	Clé	Col	Jea	Jor	Med	Rém	Rom	Sam	Yas			
T1	28,57	29,63	43,24	31,34	56,41	10	40,63	41,56	39,29	54	23,53	36,20	177,22	[10-56,41]
T2	36,21	23,60	44	37,23	27,40	42,19	31,37	32,26	20	36,11	26,42	32,43	57,02	[20-44]

Tableau 5 - Évolution des pourcentages de variantes standard pour les 3 variables

Conclusion

Les résultats que nous avons présentés ont montré deux faits importants. Premièrement, au niveau macrosociologique, notre étude a permis de montrer que la stratification sociale de l'usage des liaisons facultatives, avérée chez l'adulte, se manifeste chez l'enfant vers 5-6 ans. Deuxièmement, les deux suivis longitudinaux de groupes d'enfants ont montré que leurs usages de variantes sociolinguistiques convergent après un certain temps de fréquentation quotidienne. Comment rendre compte alors de ces effets, apparemment contradictoires, de divergence globale et de convergence locale ? Nous proposerons deux explications.

La première est d'ordre méthodologique. En effet, les grandes enquêtes sont en général menées dans des situations contrôlées tout comme notre étude expérimentale macrosociologique. Sachant que le comportement linguistique des locuteurs varie en fonction du contexte de l'échange et que cette sensibilité au contexte est socialement différenciée, notre démarche méthodologique a peut-être pu faire émerger la divergence observée dans les usages. En effet, il a été montré qu'en situation formelle, les enfants de milieux favorisés modifient davantage leurs productions en utilisant plus de variantes standard que les enfants de milieux moins favorisés (Lafontaine, 1986). Notre étude expérimentale s'apparentant à une situation formelle, les différences sociales observées pourraient être le fait de différences d'adaptation à la situation. Les enquêtes microsociologiques permettent au contraire d'appréhender les usages des locuteurs dans des situations qui leur sont familières et au sein de leur groupe de pairs, comme ce fut le cas dans nos deux études microsociologiques ; ce type de situation est alors plus propice à l'émergence d'une convergence dans les usages.

La seconde explication que nous avancerons est d'ordre sociologique. Notre idée est que, dans des groupes d'enfants qui se fréquentent quotidiennement pendant une année, les usages linguistiques convergent en fonction de la composition sociale du groupe et des influences mutuelles des individus entre eux. Les différents niveaux de convergence entre les différents groupes conduisent alors à des divergences au niveau macrosociologique. Cette hypothèse mériterait que l'on s'y attarde en observant différents groupes d'enfants de différentes écoles afin de vérifier la concomitance de la stratification macrosociologique et des convergences microsociologiques.

Références

- ÅGREN, John (1973). *Étude sur quelques liaisons facultatives dans le français de conversation radiophonique*. Uppsala : Uppsala University Press.
- ASHBY, William (1981). French Liaison as a Sociolinguistic Phenomenon, in *Linguistics Symposium on Romance Languages (9th)*, Cressey, William W & Napoli, Donna Jo (Eds). Washington, DC : Georgetown University Press, 46-57.
- BARBU, Stéphanie (2000). *Les relations sociales entre enfants à l'école maternelle : bases comportementales et langagières*. Thèse de doctorat, Université de Rennes 1.
- BOOIJ, Geert & DE JONG, Daan (1987). The domain of liaison : theories and data, *Linguistics*, 25, (5): 1005-1025.
- CHEVROT, Jean-Pierre, BEAUD, Laurence & VARGA, Renata (2000). Developmental data on a French sociolinguistic variable : the word-final post-consonantal /R/. *Language Variation and Change*, 12, (3): 295-319.
- DE JONG, Daan (1991). La liaison à Orléans (France) et à Montréal (Québec). *Mimeo of paper presented at the XIIIth international congress of phonetic sciences*. Aix-en-Provence.
- DE JONG, Daan (1994). La sociophonologie de la liaison orléanaise, in *French Generative Phonology : retrospective and perspectives*, Lyche, Chantal (Eds). Salford : Association for French Language Studies in association with the European Studies Research Institute, 95-130
- DELATTRE, Pierre (1966). *Studies in French and comparative phonetics: selected papers in French and English*. The Hague, London, Paris : Mouton.
- DESROSIÈRES, Alain & THÉVENOT, Laurent (1988). *Les catégories socioprofessionnelles*. Paris : La Découverte.
- GREEN, John N. & HINTZE, Marie-Anne (1990). Variation and Change in French linking phenomena, in *Variation and Change in French : essays presented to Rebecca Posner on the occasion of her sixtieth birthday*, Green John N. & Ayres-Bennet Wendy (Eds). London, New York : Routledge, 61-88.
- LABOV, William (1976). *Sociolinguistique*. Paris : Editions de Minuit.
- LAFONTAINE, Dominique (1986). *Le Parti pris des mots*. Bruxelles: Pierre Mardagua.
- LAKS, Bernard (1977). Contribution empirique à l'analyse socio-différentielle de la chute de /r/ dans les groupes consonantiques finals, *Langue Française*, 34: 109-125.
- LAKS, Bernard (1980). *Différenciation linguistique et différenciation sociale : quelques problèmes de sociolinguistique française*. Thèse de 3ème cycle de linguistique, Université de Paris VIII - Vincennes.
- MALÉCOT, André (1975). French liaison as a function of grammatical, phonetic and paralinguistic variables. *Phonetica*, 32: 61-179.
- MILROY, Lesley (1987). *Language and social networks*. Oxford : Blackwell.
- MOISSET, Christine (2000). *Variable liaison in Parisian French*. PhD thesis, University of Pennsylvania.
- NARDY, Aurélie (en préparation). *Acquisition des dialectes sociaux et des usages académiques de l'oral entre 2 et 5 ans : naissance du sentiment normatif et sociabilité*. Thèse de doctorat, Université Stendhal-Grenoble 3.
- ROBERTS, Julie (1994). *Acquisition of variable rules : (-t,d) deletion and (ing) production in preschool children*. Ph.D. dissertation, University of Pennsylvania.
- ROBERTS, Julie (1997). Acquisition of variable rules : a study of (-t, d) deletion in preschool children. *Journal of Child Language*, 24: 351-372.
- RUSSELL, Joan (1982). Networks and sociolinguistic variation in an African urban setting, in *Sociolinguistic variation in speech communities*, Romaine Suzanne (Eds). London : Edward Arnold, 25-140.
- SOKAL, Robert R. & ROHLF, James F. (1981). *Biometry : the Principles and Practice of Statistics in Biological Research*. New York : W.H. Freeman and company.

LES PYGMÉES BAKA DU GABON : APPROCHE SOCIOLINGUISTIQUE

Pascale PAULIN

ppaulin@ish-lyon.cnrs.fr

Laboratoire Dynamique Du langage

Institut des Sciences de l'Homme

14 Avenue Berthelot

69007 LYON

Abstract : This article shows that in a context of precocious bilingualism, prestige may play an essential role in choosing a language as well as an identity in cases of interethnic marriages. The impact of the present-day change in lifestyle on the use of the languages involved will also be examined and a concise evaluation of the vitality of the Baka language will be presented.

Introduction

Dans cet article, je m'efforcerai, après une présentation succincte de la communauté baka et de sa langue, de définir les divers contextes d'utilisation des trois différentes langues en présence dans la région où j'effectue mes recherches en ethnolinguistique, à savoir le baka, le fang et le français. Je présenterai également les divers facteurs qui influent sur le choix des langues parlées.

Dans un second temps, je donnerai un aperçu concis d'éléments lexicaux empruntés au fang et m'attacherai ensuite à la notion de vitalité d'une langue, en faisant intervenir les différents critères de l'Unesco.

Aucune étude n'ayant encore été menée sur les Baka du Gabon, cette communauté fera ici d'abord l'objet d'une brève présentation d'ordre sociolinguistique.

1. Quelle langue pour les Baka du Gabon ?

Le présent travail est une étude basée non seulement sur des observations directes mais également sur des analyses qualitatives basées sur une dizaine d'entretiens. J'ai essayé, dans la mesure du possible, de tendre à l'exhaustivité en prenant en considération quatre critères tirés de Milroy (1992) : l'âge, le sexe, l'aire géographique (village proche versus éloigné de la ville) et la situation contextuelle (« contextual styles »). Ce n'est que lorsque le critère est pertinent qu'il sera mentionné.

1.1. La communauté baka

Les Pygmées baka sont présents au Cameroun, au Congo-Brazzaville et au Gabon. C'est au Cameroun que la communauté est la plus nombreuse puisqu'on y dénombre environ 30000 individus d'après Brisson et Boursier, chiffres datant de 1979 repris comme estimation par Kilian-Hatz en 2002.

Les Pygmées baka du Gabon sont localisés au nord du pays. Le Gabon ne compte pas moins d'une cinquantaine d'ethnies parlant chacune sa langue propre ; toutes ces langues appartiennent à la même famille bantoue à l'exception du baka qui est une langue oubanguienne (donc très différente des langues environnantes, même si les langues bantoues et les langues oubanguiennes remontent en définitive toutes au Niger-Congo). Les Baka vivent donc au milieu d'une aire bantoue. L'autre particularité de cette communauté baka est leur mode de vie nomade, traditionnellement chasseurs-cueilleurs (Bahuchet, 1989).

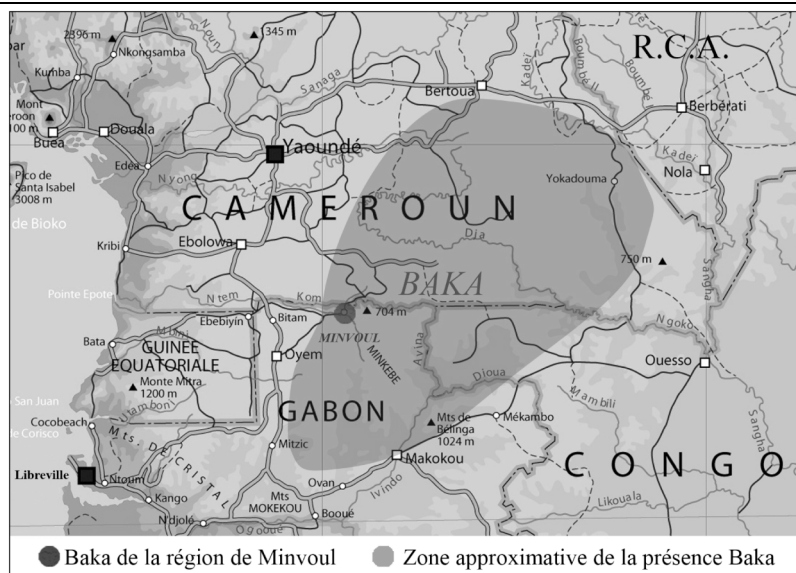


Figure 1 - Zone approximative des Pygmées baka au Cameroun, Congo-Brazzaville et Gabon

Mon travail s'attache essentiellement à la communauté baka se situant aux alentours de Minvoul, à une vingtaine de kilomètres de la frontière du Cameroun. Mvé Mebia (2001) avance le chiffre de 364 individus, localisés dans sept villages autour de la ville. Ce chiffre est approximatif car le recensement des populations traditionnellement nomades s'avère très compliqué. Il n'en demeure pas moins que les Baka sont minoritaires, et qu'ils sont entourés de la plus grande communauté, les Fang (250 000 au Gabon sur une population totale de 1 200 000 selon le recensement officiel de 1993¹) avec qui ils entretiennent des contacts réguliers. Aussi, les Baka sont bilingues dès leur plus jeune âge.

1.2. Trois langues : le baka, le fang et le français

1.2.1. Quelle langue dans quel contexte ?

Le baka est la langue maternelle de tous les individus de père et de mère baka (un seul parent baka sur les deux n'est pas suffisant pour garantir la transmission de la langue comme nous le verrons dans la partie 1.2.2.b. concernant les mariages interethniques). La transmission se fait de génération en génération, au travers notamment de toutes leurs activités quotidiennes. De même, au sein des villages baka ou lorsque la communauté se retrouve, même en ville, ils parlent leur langue entre eux.

En revanche, lorsqu'un Fang est présent les Baka, vont parler fang afin que celui-ci puisse comprendre la conversation. Cette langue est systématiquement utilisée en présence de locuteurs fang : c'est la langue de communication par excellence.

Lorsqu'ils vont en ville, ils parlent fang au marché, et même dans les commerces tenus par les Hausa², la langue utilisée est le fang.

En revanche, avec les personnes qui ne parlent pas fang, le français, qui a le statut de langue officielle au Gabon, prendra le relais. Néanmoins, tous les Baka ne parlent pas français et il

¹ Non publié.

² Une grande communauté hausa est présente au Gabon. Dans la région de Minvoul, ce sont les Hausa qui tiennent les commerces de « grande envergure » proposant une grande diversité de produits alimentaires d'importation (riz, pâtes, conserves, etc.).

est important ici de faire une distinction entre les villages proches de la ville où les habitants entendent fréquemment parler français, et les villages très retirés où le français est absent, et seulement entendu par les habitants qui se rendent en ville.

Le français est omniprésent dans tous les établissements scolaires³. La politique du gouvernement est claire quant à la langue de l'enseignement qui est essentiellement le français. Il existe de rares cas d'options de langues dites nationales⁴, comme la méthode « Rapidolangue »⁵ dans quelques établissements tenus par des missionnaires.

La situation exposée ici est en réalité bien plus complexe encore dans la mesure où dans le contexte donné précédemment (1.2.1.a.) l'emploi du baka n'est pas systématique car plusieurs facteurs influent sur le choix des langues effectivement parlées.

1.2.2. Facteurs importants influençant le choix des langues parlées

a. La notion de prestige : grande différence de statut social

Les Baka sont souvent très mal considérés par les Fang. Ces derniers ne se considèrent pas simplement comme les patrons de leurs « Pygmées » comme ils disent – ils font en effet souvent appel à eux pour défricher leurs champs – mais bien comme les propriétaires de ceux-ci. Même si, derrière cette notion, nous pouvons également voir une idée de responsabilité vis-à-vis des Baka, il n'en demeure pas moins qu'ils considèrent les Pygmées comme leurs propriétés, au même titre que leurs biens. Quelques récits nous ont même laissé entendre qu'ils n'étaient pas des êtres humains mais plutôt des animaux⁶.

Les Fang partent du principe que des personnes non scolarisées, donc de fait non éduquées, voire non « civilisées », n'ont pas grande valeur et l'ont bien fait comprendre aux Baka, qui sont persuadés d'être des créatures de rang inférieur par rapport aux Fang⁷, ces derniers étant considérés par les Baka comme inférieurs aux Blancs. Il existe donc une hiérarchie sociale où les Blancs sont en haut de l'échelle, suivis par les Fang (ou les Grands Noirs en général, par opposition aux Pygmées comme « êtres de petite taille ») et enfin par les Baka. La majorité des Baka, ainsi persuadés de leur infériorité, seront fiers de montrer qu'ils parlent fang et chercheront par tous les moyens à accéder à un statut social supérieur, notamment par le biais d'un mariage interethnique.

Vu le manque de considération que portent les Fang aux Baka, on peut se demander comment les mariages interethniques sont possibles.

³ Les établissements scolaires sont essentiellement situés dans la ville de Minvoul, aucun des villages concernant cette étude ne bénéficie d'école en son sein.

⁴ Il s'agit d'une dizaine de langues, chacune représentative d'un groupe géolinguistique différent. Par exemple, le tsgo pour le groupe B30, le punu pour le B40 et le nzébi pour le B50 (Mouguiama Daouda). Ce système de référence composé d'une lettre et d'un chiffre a été élaboré par Guthrie (1948). La lettre renvoie à une zone géolinguistique, le décimal du chiffre renvoie au groupe linguistique et l'unité à une langue particulière.

⁵ Fondation RAPONDA WALKER, pour plus de détails consulter : http://www.cyberquebec.ca/_layout/?uri=http://www.cyberquebec.ca/rapondawalker/.

⁶ « Si, ce sont les Pygmées de ma grand-mère, ils lui appartiennent. Et de toutes manières, ils ne sont pas vraiment des Hommes, ils sont comme les animaux. » Entretien avec S.P., informateur, Docteur en Philosophie, Lyon, Laboratoire DDL, novembre 2004.

« Ce ne sont pas des Hommes, ce sont des animaux pour arriver à vivre comme ça dans la forêt. » Entretien avec C.T., Makokou, informateur, février 2007.

⁷ La situation est présentée de la sorte car les arguments pris en considération concernent généralement les activités des Fang, comme la scolarité ou l'agriculture, les compétences spécifiques aux Baka ne sont jamais mentionnées, comme la chasse ou la pharmacopée.

b. Les mariages interethniques : deux langues en concurrence ?

Au Gabon, un homme désirant épouser une femme se doit de rassembler une dot, en guise de compensation de la force de production perdue dans la famille de la mariée. Or, il s'avère que la dot chez les Fang est la plus élevée des ethnies du Gabon d'après Mayer. Il explique cela par le fait que les Fang sont patrilinéaires c'est-à-dire que les enfants appartiennent à la lignée du père, et pratiquent la patrilocalité, la femme partant vivre dans la communauté de son mari. Ainsi les parents de la jeune fille perdent doublement les forces de production.

« Dans les ethnies patrilinéaires, la dot compense théoriquement deux types de "perte", l'épouse en tant que génitrice, et l'épouse en tant que productrice agricole. » (Mayer, 2002 :196).

Aussi est-il très difficile pour un jeune homme de rassembler une telle dot ; s'il est désireux de se marier rapidement, il va se « rabattre » en quelque sorte sur une épouse baka pour laquelle la dot est généralement dix fois moins conséquente que chez les Fang. L'homme fang y trouve son compte mais également la femme baka qui accèdera à un statut social dit « supérieur ». Cela est d'autant plus vrai pour ses enfants. Les Baka et les Fang étant deux sociétés patrilinéaires, les enfants d'un tel mariage seront considérés comme Fang. La mère ne voit alors pas du tout l'intérêt de parler baka à ses enfants, et souvent, dans 50% des cas, nous avons constaté que le père interdisait à la femme de parler sa langue maternelle à ses enfants. Ces derniers n'ont, dans ce cas précis, aucun patronyme baka mais seulement un patronyme fang et chrétien. Ces enfants sont des locuteurs perdus pour la langue baka, et même s'ils comprennent les conversations courantes et peuvent s'exprimer en baka (la mère continue à avoir des contacts avec sa communauté d'origine et les enfants suivent régulièrement leur mère), ils ne le font pas car ils marquent ainsi leur appartenance à la lignée supérieure des Fang. Aussi, il est hors de question, pour les enfants issus d'un tel mariage, d'envisager une union avec un ou une Baka (les jeunes hommes sont d'ailleurs plus virulents que les jeunes femmes à ce propos).

Les mariages interethniques sont de plus en plus nombreux, chacune des parties y trouvant son compte. Et ils se réalisent de plus en plus tôt, c'est-à-dire que selon la coutume du mariage enfantin⁸, une fillette baka de 3 ou 4 ans pourra être emmenée dans la famille de son futur époux afin d'y être éduquée. Elle est ainsi sortie très tôt non seulement de son environnement linguistique mais également des diverses pratiques de sa communauté. Ce qui nous amène au point suivant.

c. Le changement de mode de vie : les nomades se sédentarisent...

Les Baka sont traditionnellement des chasseurs-cueilleurs, ils puisent encore la majeure partie de leur subsistance dans la forêt. Anciennement, ils étaient nomades, mais la politique gouvernementale de sédentarisation les oblige à construire des maisons en terre – similaires à celles des Fang – et leur interdit leur hutte traditionnelle. Toutes ces directives ne sont pas suivies, mais le village est tout de même devenu le lieu de retour systématique des déplacements en forêt. Les Baka passent ainsi moins de temps en forêt et installent quelques plantations autour du village. Cette activité est relativement récente, puisqu'elle date de moins d'un siècle.

Comme cela avait déjà été noté chez Bahuchet (1989), les villages pygmées jouxtent ceux des Fang. En effet, les Grands Noirs font souvent appel aux Pygmées pour diverses tâches (le défrichage, la chasse...). Les Baka tendent à vivre comme les villageois et aspirent au confort matériel. Ils se rendent ainsi de plus en plus dépendants de travaux rémunérés, et

⁸ MAYER, Raymond, Professeur en Anthropologie, Communication personnelle, Lyon, DDL, mai 2005.

passent alors plus de temps en compagnie des Fang, parlant la langue de leurs voisins aux dépens de la leur⁹.

d. La scolarisation : la langue baka n'a pas sa place à l'école

Avant de parler de la scolarisation en tant que telle, il est important d'évoquer l'accès à l'enseignement car le facteur d'éloignement est un problème crucial. Hormis le village de Mféfélam (ou Essang), les six autres villages pygmées se trouvent à 8 kilomètres (ou plus) des établissements scolaires.

« Le 10 juillet 2002, il a été porté à notre connaissance, un cas de violation de droits humains relatif au Droit à l'éducation des enfants. [...] “ Doum-Assi ”, à quelques 8 km de Minvoul, dans la province du Woleu-Ntem, un groupe d'enfants Pygmées non-scolarisés a manifesté auprès de certains citoyens, leur volonté de voir l'Etat garantir leur scolarisation. » (Mba Abessole, 2004 : 194)

A ce jour, les enfants de Doumassi, comme les autres villages éloignés, n'ont toujours pas accès à un enseignement de proximité.

Lorsqu'ils vont à l'école, quelle langue les enfants pygmées utilisent-ils ? Comme je l'ai évoqué précédemment, les Baka sont victimes de dépréciations évidentes. Ainsi pour ne pas se faire remarquer, ils préfèrent parler fang plutôt que baka dans la cour de récréation, et en classe, la seule langue acceptée est le français. Les enfants baka se font quotidiennement railler par leurs camarades fang. Aussi, bon nombre d'entre eux refusent-ils d'achever l'année commencée. Et même si certains parents voient la nécessité d'envoyer leur enfant à l'école, entre autres, afin d'arriver à faire valoir leurs droits aux yeux des diverses autorités du pays, ils n'arrivent pas à forcer l'enfant à poursuivre sa scolarité¹⁰. Situation qui peut néanmoins profiter à la préservation de la langue baka qui n'a pas sa place à l'école, et qui nous amène au point suivant concernant la vitalité de la langue baka.

2. Le baka : une langue en danger ?

J'aborderais ici deux points spécifiques : celui des emprunts au fang puis l'évaluation de la vitalité de la langue baka.

2.1. Les emprunts non repérés

Je ne parlerai pas ici des emprunts facilement repérables par les locuteurs eux-mêmes, c'est-à-dire des mots identiques ou quasi identiques à ceux trouvés en fang¹¹, mais des emprunts anciens au fang qui ont été bien intégrés en baka. Les locuteurs ont alors plus de mal à repérer les indices de l'emprunt, ils sont persuadés qu'il ne s'agit pas d'un emprunt à la langue voisine et ont de fait des difficultés à fournir le terme dans leur langue propre. Nous rencontrons des attestations de ce cas de figure dans le lexique de base où il n'y a pas eu de néologisme interne mais des néologismes externes, c'est-à-dire des emprunts avec

⁹ Mon propos n'est pas ici d'obliger les Baka à continuer de parler leur langue, mais de constater que, du point de vue de la diversité linguistique, la disparition d'une langue est dommageable. La diversité – expression de la complexité – intéresse la science et il est évident que l'étude d'une langue met en lumière des phénomènes propres à la langue mais éclaire également d'autres champs d'investigations scientifiques (l'histoire, l'anthropologie, l'archéologie, la génétique, etc.).

¹⁰ D'autant plus que le rythme scolaire, pensé pour des populations sédentaires, ne convient nullement aux Baka qui effectuent tout au long de l'année de multiples déplacements, souvent de longue durée en pleine forêt, pour de grandes chasses et ainsi transmettre une partie de leurs savoir-faire ou simplement pour renforcer l'unité familiale.

¹¹ Il s'agit d'emprunts récents.

appropriation. La situation s'avère préoccupante lorsqu'il s'agit des lexiques spécialisés¹² comme ceux de la faune ou de la flore en rapport étroit avec l'environnement proche et quotidien des Baka. Certains termes ont été empruntés au fang alors qu'il existait un mot en baka. Il est très étonnant dans ce cas d'avoir des emprunts fang alors que les Baka sont beaucoup plus souvent en contact avec ces espèces, botaniques par exemple. Néanmoins, il paraît difficile que des termes liés à des pratiques spécifiques baka, n'ayant aucun équivalent en fang, disparaissent, comme par exemple les différents termes dénotant le miel. J'ai collecté huit termes correspondant à huit variétés différentes de miel, or il semblerait que les Fang n'en dénomment que quatre, même si par ailleurs, ils peuvent connaître plus de quatre variétés¹³.

En revanche, on peut envisager qu'ils adoptent la langue du voisin, c'est-à-dire qu'ils ne parlent finalement plus que le fang et que les termes spécifiques baka se maintiennent dans le lexique. Il se peut également que l'on ait affaire à une série d'hyperonymes c'est-à-dire qu'un terme baka se généralise et désigne l'ensemble des éléments d'une catégorie, comme par exemple « ija » le terme générique pour « éléphant » serait utilisé pour tous les éléphants quelque soit leur taille, leur âge ou leur sexe, alors qu'il existe six termes pour des éléphants mâles de taille différente, un terme pour l'éléphante et un terme pour l'éléphanteau.

Mais il ne faut pas non plus se leurrer car si leur mode de vie change, leurs pratiques spécifiques tendront à s'estomper, voire à disparaître totalement. Nous sommes en droit de nous demander si la langue baka résistera à toutes ces pressions et si elle est réellement en danger. Ce qui nous amène au point suivant, qui concerne les critères que l'Unesco prône pour l'évaluation de la vitalité des langues. Chaque critère comporte un degré. Ces degrés sont notés de 0 à 5 et si l'un des critères est au degré « 0 », il y aurait une menace. Dans ce cas, les autres critères doivent impérativement atteindre un bon degré sinon la disparition de la langue est à craindre.

2.2. *Evaluation de la vitalité de la langue baka*

Critères de l'Unesco : langue baka	Degré
1. Transmission intergénérationnelle	5 ou 4
2. Nombre de locuteurs	± 500
3. Proportion de locuteurs au sein de la communauté	4
4. Extension ou régression des domaines d'utilisation de la langue	3 ou 4
5. Réponse de la langue face aux nouveaux domaines et dans les médias	0
6. Matériel disponible pour l'enseignement et la littérature	0
7. Attitude des Institutions et du Gouvernement face à la langue	1
8. Attitude des locuteurs envers leur propre langue	1
9. Qualité et type de documentation disponible	2 ou 3

Tableau 1 - La langue baka face aux neuf critères mis en place par l'Unesco (2003)

L'application de ces divers critères¹⁴ n'est pas toujours simple, mais ils ont le mérite de poser un cadre, et de fait d'induire des discussions autour de la notion de vitalité de la langue¹⁵.

¹² Pour une discussion plus approfondie sur la résistance à l'emprunt des différents types de lexiques voir Hock & Joseph (1996) pp. 274-275.

¹³ Ce travail de terrain concernant les dénominations du miel en fang est en cours mais il semble peu probable pour Piher Médjo Mvé – Maître de Conférences en Linguistique – d'arriver à trouver des personnes capables de proposer huit termes différents. Communication personnelle, Lyon, DDL, décembre 2006.

¹⁴ Chaque degré est spécifié en fonction du critère. Prenons l'exemple du 3^{ème} facteur concernant la proportion de locuteurs au sein de la communauté, en résumé : « Tous les membres de la communauté parlent la langue. » renvoie au degré 5, « Presque tous parlent la langue. » degré 4, « Une majorité parle la langue. » degré 3, « Une minorité parle la langue. » degré 2, « Très peu parlent la langue. » degré 1 et enfin « Personne ne parle plus la langue. » degré 0 (UNESCO).

Pour ce qui concerne le premier critère, la transmission intergénérationnelle, comme je l'ai évoqué précédemment, la transmission de la langue de génération en génération est effective. Néanmoins, elle n'atteint pas le degré maximum si nous prenons en considération les mariages interethniques. Le même argument prévaut pour le troisième critère, la proportion de locuteurs au sein de la communauté, qui, de ce fait, ne peut atteindre le degré 5. Revenons au second critère concernant le nombre de locuteurs peu élevé au Gabon. La communauté baka se retrouve de plus en plus isolée du fait de la politique de sédentarisation qui implique une diminution des grands déplacements de population, même vers le Cameroun où les Baka, bien plus nombreux, pourraient jouer un rôle de « revitalisation » linguistique, voire culturelle. De plus, le baka au Gabon est la seule langue n'appartenant pas au grand ensemble des langues bantoues, et nous avons l'exemple de toutes les autres communautés pygmées présentes au Gabon qui ne parlent plus que la langue de leurs voisins. Le quatrième critère, l'extension ou la régression des domaines d'utilisation de la langue, semble atteindre le degré 4 concernant la parité multilingue dans le sens où chaque langue aurait son champ d'action, mais la réalité offre un tableau moins prometteur si l'on prend en compte le changement de mode de vie en cours évoqué dans la première partie. Aussi, les nouvelles activités auxquelles auront accès les Baka seront-elles communes aux Fang et la solution du recours à l'emprunt lexical paraît la plus économique et la plus évidente, surtout en cas de bilinguisme. Le critère suivant concernant la réponse de la langue face aux nouveaux domaines et dans les médias est clairement déficient. Aucune langue gabonaise n'a été instrumentalisée (au sens où on l'entend pour le swahili par exemple), et même les dix langues dites nationales ont des difficultés à se faire une place dans les médias : on imagine mal comment le baka pourrait y avoir sa place¹⁵. Quant au matériel disponible pour l'enseignement, il n'y en a aucun. Quelques écrits portant sur le baka existent : des dictionnaires et des contes. Ceux-ci sont de très bonne qualité, mais n'ont pas du tout vocation à servir l'enseignement de la langue en question. De plus, la population n'a aucunement accès à ces documents, d'autant que la plupart d'entre eux ne sont pas publiés. Le septième critère au sujet de l'attitude des Institutions et du Gouvernement face à la langue fait clairement défaut. Aucune réelle politique de préservation ne serait-ce que de la dizaine de langues nationales n'est mise en place par le gouvernement gabonais. Rien n'est fait pour encourager l'emploi des différentes langues présentes dans le pays. Néanmoins, l'utilisation de ces diverses langues n'est pas prohibée : c'est pourquoi nous avons affaire ici au degré 1. Le point suivant concernant l'attitude des locuteurs envers leur propre langue me paraît essentiel pour la communauté que j'étudie. Prenons un exemple bien concret, qui peut sembler anecdotique, mais qui prend toute sa signification vu la situation présentée précédemment. Lors de mon troisième séjour chez les Baka, un jeune membre de la communauté est décédé le lendemain de mon arrivée. Il habitait le village le plus proche de la ville de Minvoul – Mfêfélam, à peine deux kilomètres de distance – village où les contacts entre les Fang et les Baka sont quotidiens, quasi-permanents, à tel enseigne que le village dispose de sa propre église. Le défunt fut enterré à la méthode fang et le discours fut prononcé en fang par un parent. Or, un autre parent, d'un village très éloigné de la ville, où les contacts avec les Fang ne sont pratiquement qu'à l'initiative des Baka, s'est insurgé en disant que le défunt était baka et que le discours devait être prononcé en langue baka, ce qu'il

¹⁵ Prenons l'exemple du 3^{ème} critère résumé dans la note précédente : qui sont les membres qui composent la « communauté » ? Devons nous prendre en considération les enfants issus de mariage mixte ? Ce sont deux sociétés patrilineaires, et malgré la paternité reconnue par un Fang, les Baka considèrent ces enfants comme des membres de leur communauté. Doit-on privilégier le point de vue endogène ou le point de vue exogène ?

¹⁶ Il existe un programme d'éducation populaire à la télévision et à la radio réalisé dans les 10 langues « nationales » (Mouguiama Daouda).

fit aussitôt. Nous voyons ici que l'attitude du locuteur est très importante quant à la survie de sa langue et malgré la notion de prestige évoquée dans la première partie, certaines personnes refusent de voir leur langue disparaître et à travers elle certains traits culturels. Le dernier critère, la qualité et type de documentation disponible, rejoint la question du matériel disponible car même s'il existe des documents de très bonne qualité, il est impossible pour la communauté de se les procurer. De plus, les documents en question ne concernent que la (ou les) variété(s) de langue baka du Cameroun ; aucun document concernant le baka du Gabon n'existe pour le moment. C'est d'ailleurs une des raisons pour lesquelles je me suis intéressée à cette communauté.

Conclusion

Comme nous l'avons vu, le baka, le fang et le français sont trois langues qui ne jouissent pas d'un statut identique, que ce soit au regard du Gouvernement ou aux yeux de la communauté baka. La langue de ces derniers a tendance à être délaissée au profit du fang en raison de la différence de prestige entre les deux communautés en présence. Et l'un des facteurs déterminants dans la préservation de la langue est le changement de leur mode de vie : plus le changement sera brutal plus le risque de disparition de la langue est à craindre. Aussi me semble-t-il essentiel de prendre en considération les spécificités de cette communauté, ne serait-ce que dans le système scolaire où l'enseignement n'est absolument pas adapté à leur mode de vie encore en partie nomade. De plus, une politique de valorisation des diverses pratiques spécifiques des Baka tendrait à réduire, voire à supprimer, cette différence de statut social liée à la notion de prestige et par là même à favoriser la préservation de leur langue. L'analyse des différents critères de l'Unesco évoqués précédemment ne fait que renforcer l'idée que la langue est effectivement en danger. La situation est d'autant plus préoccupante que les locuteurs eux-mêmes ne sont pas toujours conscients de cette situation, ni des enjeux, et surtout qu'ils ne voient pas réellement l'intérêt de la préservation de leur langue, excepté pour certaines pratiques où le secret est de mise, comme par exemple les rites initiatiques.

Bibliographie

- BAHUCHET, Serge (1989). *Les Pygmées aka et baka, Contribution de l'ethnolinguistique à l'histoire des populations forestières d'Afrique Centrale*. Thèse de Doctorat d'Etat. Paris V : Université René Descartes.
- BRISSON, Robert & BOURSIER, Daniel (1979). *Petit dictionnaire baka - français*. Douala : non publié.
- GUTHRIE, Malcolm (1948). *The Classification of the Bantu Languages*. Londres : IAI.
- HOCK, Hans Henrich & JOSEPH, Brian D. (1996). *Language History, Language Change, and Language Relationship*. Berlin & New York : Mouton de Gruyter.
- KILIAN-HATZ, Christa (2002). *Contes des Bakas (Cameroun)*. Cologne : Rüdiger Köppe Verlag.
- MAYER, Raymond (2002). *Histoire de la famille gabonaise*. Libreville : Editions du LUTO.
- MBA ABESOLE, Paul (2004). *Le livre blanc des droits humains au Gabon*, Libreville : Ministère des Droits de l'Homme, Multipress.
- MVE MEBIA, Emmanuel (2001). *Habitat baka et organisation socio-économique dans la périphérie nord et nord-ouest de la réserve de Minkébé (Gabon)*, DESS Tourisme Culturel, Université Omar Bongo, Libreville.
- MILROY, James (1992). *Linguistic Variation & Change*. Oxford & Cambridge : Blackwell.
- MOUGUAMA DAOUDA, Patrick, Professeur en Linguistique, Communication personnelle, Lyon, DDL, 2005.
- UNESCO, Intangible Cultural Heritage Section's Ad Hoc Expert Group on Endangered Languages (2003). *Language Vitality and Endangerment*, Paris.

CONTACTS DES LANGUES ET IDENTITÉ : LE RÔLE DE L'INNOVATION LEXICALE DANS LE CAS DU SCAMTO

Lerato Caroline POOE

leratop@hotmail.com

UMR 6039 Bases, Corpus et Langage

Université de Nice-Sophia Antipolis

06204 Nice cedex

Abstract : Many authors highlight important imperatives in relation to the « deliberate and conscious marking of identity » in the emergence of mixed languages. From this perspective, and in the particular case of Scamto (a South African Township mixed code), this article focuses on the question of lexical manipulation considered as a conscious process of lexical creation through which its creators and speakers express their new identity.

Introduction

Le scamto, aussi connu sous les noms *tsotsitaal*, *flaaitaal*, *isicamtho*, *stsotsi sjita*, *spantsula*, *lingo* etc. est une langue émergente de la région de Soweto. Il est issu du contact entre différentes langues sud-africaines (langues bantoues : zoulou, sotho septentrional et méridional, xhosa, swati, tswana, venda, ndébélé, tsonga, et langues germaniques : anglais et afrikaans). C'est un parler de nature mixte qui reflète le contexte multilingue des townships noirs en Afrique du Sud dans lequel il se développe. La structure morphosyntaxique du scamto est bantoue, venant surtout du sotho et du zoulou, mais son lexique vient de différentes sources (anglais, afrikaans, xhosa, zoulou, sotho, gestes, etc.), et c'est cette hétérogénéité qui souligne la différence entre le scamto et les langues sources de sa grammaire.

Le scamto pris comme objet d'étude nous permet de nous interroger sur le rôle de la « volonté » des locuteurs dans la formation de certains parlers comme les langues mixtes/mélangées ou les langues de jeunes. Pourquoi et comment les nouvelles façons de parler naissent-elles ? Les travaux sur les pidgins ont montré que ceux-ci résultent d'un besoin de communiquer dans des situations précises et limitées, mais cela n'est pas pertinent pour les langues mixtes et de nombreux auteurs mettent en avant des impératifs autres que le besoin d'une langue commune dans la formation de ces langues, impératifs qui dans de nombreux cas sont en rapport avec « un marquage conscient et délibéré d'identité ». (cf. entre autres : Le Page & Tabouret-Keller, 1985 ; Muysken, 1997 ; Golovko, 2003). Dans cette perspective, et dans le cas particulier du scamto, on s'intéresse à la question de la « réorientation lexicale ». Comment l'analyser ? Que manifeste-t-elle ? Avons-nous affaire ici à un cas de manipulation délibérée de langue (Golovko, 2003 ; Mous, 2003) ? C'est à l'étude de ce procès de « réorientation lexicale », de *manipulation lexicale* que Mous considère comme un processus conscient de l'innovation/de la création lexicale, que nous nous attachons.

L'étude prend appui sur une série d'enquêtes menées dans la Province de Gauteng, notamment à Soweto. Deux types de données recueillies ont été utilisés : un **échantillon lexical** du scamto extrait des conversations spontanées enregistrées dans des townships de Soweto et **des représentations** des locuteurs/usagers et non-usagers (partageant le même espace anthropologique que les usagers) extraites des discussions (enregistrées) portant sur l'émergence, la nature, l'emploi, le statut etc. du scamto. Le traitement consiste tout d'abord

à montrer la pertinence de la manipulation lexicale pour le scamto à travers les stratégies mises en œuvre dans l'innovation lexicale, puis à apprécier la signification identitaire de ces processus/stratégies en prenant également en compte les éléments identitaires exprimés à un autre niveau dans les représentations que les locuteurs et les non-locuteurs se font du scamto.

En nous interrogeant sur la pertinence de ces notions (marquage conscient d'une identité, et la manipulation lexicale) pour le scamto, nous poserons deux questions principales : quels processus (stratégies) de manipulation lexicale sont les plus importants dans le cas du scamto ? Jusqu'à quel point la « manipulation lexicale » peut-elle révéler des éléments identitaires des locuteurs ?

1. Les townships de Soweto

Il convient tout d'abord de présenter le contexte historique, politique et social dans lequel ce parler évolue, celui des townships de Soweto. Soweto est l'acronyme de « *South-West Townships* ». Situé à 20 Km au sud-est de la ville de Johannesburg, c'est l'ensemble de vingt-six townships plus ou moins séparés et qui couvre environ 40 miles. Avec plus de quatre millions d'habitants (recensement de 1996) c'est la plus grande communauté noire en Afrique du Sud, produit dès l'origine, d'un aménagement ségrégationniste. La ville de Johannesburg et ces townships de Soweto se sont développés grâce à la découverte de l'or dans la région. Cette découverte en 1886 a entraîné une explosion de la migration. Soweto était à l'époque une vaste plaine. Les mines et l'élargissement de la ville de Johannesburg avaient besoin d'une main-d'œuvre bon marché et les Africains ayant besoin de travailler sont venus dans cette région. C'était le début de la séparation raciale qui a été continuée plus tard par le régime de l'apartheid. L'histoire de Soweto est riche concernant la lutte contre cette politique.

Les débuts de Soweto datent de 1904 et ses différentes communautés ont toujours été caractérisées par la multiethnicité, l'hétérogénéité, une forte interaction sociale et l'émergence des *sub-cultures*. Quand la politique de la ségrégation a été mise en place, la ville fut réservée à l'habitation des Blancs. À Soweto la répartition du territoire s'est faite selon les différentes ethnies et ainsi selon les langues. Ces différents townships marquent donc les différents processus de la dépossession, de la distribution du terrain et de l'urbanisation pendant les années de la discrimination raciale en Afrique du Sud. Très vite, Soweto est devenu le fief de la contestation politique africaine : émergence des grands leaders noirs, futurs cadres du pouvoir noir comme Mandela, Sisulu etc. ; campagnes pour l'amélioration des conditions de vie, squat de masse ; rapprochement avec les partis de la gauche blanche opposée à la ségrégation. Soweto a connu un accroissement démographique fulgurant mais imprévu marqué par la construction de bidonvilles et l'abandon quasi-total des services publics. Malgré les efforts du gouvernement pour contrôler l'arrivée des travailleurs noirs dans les villes, beaucoup sont venus de la campagne et des pays voisins pour chercher du travail dans la ville d'or. En 1976, la révolte finale est venue de la jeunesse et le détonateur fut l'imposition de l'enseignement obligatoire en afrikaans. Les revendications concernaient plus que la suppression de l'afrikaans. Les Noirs des townships se révoltaient contre le régime de l'apartheid qui les restreignait dans leur liberté et les contraignait à une éducation médiocre pour qu'ils restent au plus bas de l'échelle sociale. Durant les années 80, le township fut le symbole de la résistance noire à l'apartheid. Les problèmes à Soweto ont toujours été le logement, le surpeuplement, le chômage et une infrastructure pauvre qui ont entraîné d'autres problèmes de délinquance avec l'émergence de nombreux gangs, problèmes qui font aujourd'hui la mauvaise réputation de Soweto et de Johannesburg. C'est seulement depuis le nouveau régime démocratique que le gouvernement essaye d'améliorer les conditions de vie dans les townships. C'est un lieu de contrastes qui

fait se côtoyer d'un côté des quartiers bourgeois et recherchés, de l'autre des quartiers modestes puis des bidonvilles de la misère ; la jeunesse délinquante et les jeunes diplômés.

Aujourd'hui Soweto est un *melting pot* des cultures sud-africaines et a développé ses propres *sub-cultures* surtout pour les jeunes. L'influence afro-américaine est très forte mais adaptée aux conditions locales. Dans leurs discours, leurs façons de parler, leurs modes vestimentaires et leurs comportements, les Sowetans dégagent une sophistication cosmopolite. Grâce à sa proximité à la ville de Johannesburg et à l'économie du pays, c'est le plus métropolitain de tous les townships du pays – qui lance des nouvelles modes en politique mais aussi des modes vestimentaires, musicales et artistiques (comme la danse) et aussi langagières. Sur le plan linguistique, Soweto connaît un plurilinguisme qui englobe les onze langues officielles du pays et leurs nombreuses variétés : variétés régionales, urbaines/rurales ; variétés stylistiques ainsi que d'autres langues minoritaires - une situation qui favorise des dynamiques linguistiques complexes. En effet, les pratiques langagières, phénomène complexe, par essence, sont rendues encore plus complexes dans les townships de Soweto (Ntshangase, 2002 : 407) où certaines pratiques semblent traverser toutes les barrières linguistiques, politiques et ethniques. Dans ces townships le plurilinguisme est la norme et, peut-être plus qu'ailleurs, les habitants ont à leur disposition ce que Nicolaï (2003 : 82) appelle un « *répertoire non-fini* » de modes de s'exprimer. Les modes d'expressions s'entremêlent, les langues sont alternées voire mélangées dans les discours ordinaires. Et certaines pratiques mènent à la création de nouveaux parlars comme le scamto. En effet, nous posons comme hypothèse principale que le répertoire des codes des individus vivant dans les townships de Soweto est « un espace non-fini de recomposition linguistique continue dans lequel se manifeste un *feuilletage* inhérent à la fois à la fonctionnalisation des langues et à celle du langage. » (Nicolaï, 2003 : 82). Ainsi, le phénomène que l'on appelle scamto relève du *feuilletage* en tant que résultat d'une opération conduite par des 'acteurs' (locuteurs/créateurs). Le scamto serait alors une illustration du feuilletage car il véhicule de nouvelles normes et une nouvelle identité. Une hypothèse secondaire pourrait permettre de mieux comprendre le rapport entre le répertoire non-fini et la création d'une nouvelle identité : la *manipulation lexicale* (au sens de Mous, 2003) est un processus conscient de l'innovation lexicale par lequel les 'acteurs'/locuteurs expriment leur nouvelle identité.

2. Manipulation lexicale

Selon Mous (2003 : 217-18), la manipulation lexicale ou dans d'autres termes, la construction d'un lexique parallèle se manifeste dans diverses situations comme les registres de respect, le tabou, l'argot de groupes sociaux, les langues de jeunes, les codes secrets, les langues de rites etc. Le facteur le plus important est le fait que les gens soient conscients de la forme de lexique qu'ils utilisent. Dans le cas précis du scamto, la plupart du lexique existe en parallèle avec d'autres formes lexicales de leur répertoire non-fini. Les locuteurs ne créent ni n'utilisent le lexique du scamto à cause d'un besoin de mots pour communiquer. Leur choix est conscient, un choix de communiquer en utilisant un lexique qui exprime leur identité : comme le déclare un locuteur : « en scamto un mot existe parce qu'il dit beaucoup de choses sur les gens ». Les locuteurs se trouvent donc dans une situation où ils peuvent choisir la forme à employer, leur choix est ainsi presque toujours conscient et fonctionnel. Les locuteurs sont alors capables d'exercer un certain contrôle sur leur usage de différents modes d'expression dans leur répertoire.

Les discours épilinguistiques montrent un rapport entre les éléments identitaires révélés par la manipulation lexicale et les représentations des usagers et non-usagers qui décrivent le scamto tout à la fois comme : symbole d'une nouvelle identité de la communauté noire dans la nouvelle Afrique du Sud ; le véhicule d'une identité locale ; un symbole de la diversité en Afrique du Sud, une langue urbaine, le reflet d'une société en progression, le symbole du

développement, une langue de liberté et de modernité, une langue qui unifie le peuple sud-africain, une langue de gangsters, l'expression d'un manque de respect, une langue crue, libre, vive, jeune, anti-frontalière, anti-apartheid, etc.

3. Les stratégies de manipulation lexicale en scamto

Selon Mous (2003), la fonction de la manipulation lexicale détermine quelle stratégie est adéquate dans un parler donné. Le scamto qui sert principalement à exprimer une nouvelle identité semble préférer des stratégies que nous découvrons ci-dessous. Il s'agit d'examiner comment la manipulation lexicale se produit d'une manière consciente dans le cas du scamto.

3.1. L'hybridation

Nous appelons « hybride » un mot constitué d'éléments provenant de plusieurs langues différentes (Lafage, 1998 : 282). Représentant l'hétérogénéité du contexte linguistique des townships de Soweto, les particularités lexicales hybrides sont de diverses origines. Cependant, leur principale caractéristique commune c'est que ce sont des combinaisons des lexèmes issus principalement de l'anglais et de l'afrikaans et des affixes des langues bantoues (sotho ou zoulou). Tous les verbes provenant de l'anglais ou de l'afrikaans sont greffés d'un affixe sotho ou zoulou. Le verbe 'buy' (acheter en anglais) devient *baya* en scamto : on ajoute le suffixe *-a* au lexème anglais respectant donc la forme des verbes sotho. Au passé on aura *bayile*, ajoutant donc le suffixe *-ile* du passé. Pour les noms, le pluriel est souvent formé par l'ajout d'un préfixe (*ama-/ma-/di-* etc.) comme dans les noms bantous. Par exemple, 'demons' (anglais pour démons) devient *madimon*.

Les hybrides sont des indicateurs de la nouvelle identité véhiculée par le scamto, principalement celle de la diversité culturelle : ils témoignent d'une situation sociolinguistique de plurilinguisme, et d'une hétérogénéité linguistique qui entraîne des mélanges de langues de différentes familles, fonctionnant alors comme des indices d'homogénéisation, ce qui reflète la diversité des townships et du pays en général, qui s'harmonise progressivement en une seule identité sud-africaine. De telles stratégies montrent une grande liberté d'expression et de comportement. Il faut rappeler que le régime de l'apartheid interdisait les mélanges entre les gens de différentes races/ethnies et surtout entre les Noirs et les Blancs. Le fait de greffer des affixes des langues bantoues aux mots venant des langues germaniques constitue alors une revendication d'une liberté longtemps désirée de la part des locuteurs, et pour certains, l'affichage de leur respect.

3.2. La relexification

La plupart des mots qui comportent des sons qui ne sont pas dans le phonétisme bantou restent tels qu'ils sont en scamto. Souvent quand il s'agit des mots anglais, leur prononciation en scamto correspond à celle du BSAE¹. Cependant certains mots venant de l'anglais ou de l'afrikaans subissent des changements de sons pour les rendre conformes aux structures du scamto. C'est ce que l'on appelle dans ce travail la « relexification » (Bakker, 1997), en d'autres termes, le changement de la forme phonétique des signes lexicaux. Certains sons sont remplacés par d'autres des langues bantoues. *Huzethi* ? (ça va ?) en scamto vient de l'anglais « *how is it* ? ». Certaines relexifications entraînent aussi un changement de sens comme pour *ankele*, un mot que l'on utilise pour désigner une personne respectable mais que l'on considère comme un peu démodée, ce qui est différent du sens du mot source « *uncle* » (oncle en anglais). Le mot prend donc un sens plutôt péjoratif en scamto. Les relexifications sont un moyen de bantousation du lexique venant de l'anglais et de l'afrikaans. Les distorsions phonétiques sont un moyen important de souligner la

¹ Black South African English

différence entre la langue source et le scamto et de faire ressortir le côté africain (bantou) du parler, et ainsi le non mépris des langues africaines dans un contexte plurilingue.

3.3. *La troncation*

C'est l'une des stratégies les plus courantes de la manipulation lexicale notamment pour les langues mixtes, les langues de jeunes et les argots. En scamto nous retrouvons des troncations comme : « *bra* » venant du mot polysyllabique « brother » (frère et parfois camarade en anglais). En scamto le mot désigne surtout « ami/pote ». Souvent les troncations démontrent le côté familial du scamto ou créent une sorte de violence phonotactique qui est souvent perçue comme un manque de respect.

3.4. *Métaplasmes*

D'autres stratégies qui souvent ont le même effet que la troncation sont les métaplasmes tels que l'aphérèse : souvent un phonème est supprimé en début du mot en scamto, par exemple, *stlamatlama* (gueule de bois) au lieu de *setlamatlama* (plante) comme le mot se prononce en sotho. On constate bien sûr dans cet exemple un changement de sens. Un autre type de métaplasme courant en scamto est l'épenthèse : par exemple, pour une VW Golf III VR6, on utilise le mot *vorara* en scamto. L'origine de ce mot est le « VR » de VR6 qui s'est transformé en l'onomatopée *vrr* qui imite le bruit d'une voiture. On a ensuite ajouté des voyelles entre les consonnes pour adopter le système *cvcvcv* courant dans les langues bantoues concernées (sotho et zoulou).

3.5. *Siglaions/abréviations*

Le scamto a de nombreux sigles formés à partir des mots venant des langues sources. Par exemple, *TP* (prononciation anglaise) vient du mot anglais *toilet paper* (papier toilette). Un autre exemple est *Z3* (*z three*) venant de *Zonke zinto zimphethe* (zoulou), littéralement, 'elle a la maladie' et qui est utilisé pour le sida. *Jozi* est l'abréviation de *Johannesburg*. L'emploi de ces sigles et abréviations non seulement est interprété comme du modernisme mais « permet d'éviter de prononcer des mots qui ne sont parfois pas très agréables » déclare un locuteur.

3.6. *Allongements*

L'inverse des siglaions se produit aussi : de nombreuses abréviations ou des mots courts sont allongés en scamto. On retrouve cette stratégie souvent dans le lexique de boissons alcoolisées. Pour *J&B* (un whisky) on dit « *Jesu le Barutuoa* » en scamto, venant du sotho et littéralement, 'Jésus et disciples'. Pour une bière sud-africaine assez populaire 'Castle' on dit, *Children Attend School To Learn English*, littéralement, 'les enfants vont à l'école pour apprendre l'anglais' ou également, *Can Adam Still Tickle Little Eve?* littéralement 'Adam peut-il toujours chatouiller Eve?' Nous constatons dans certaines de ces appellations des connotations religieuses ou sexuelles – utiliser des éléments religieux dans des expressions populaires peut être amusant pour certains individus mais provocant pour d'autres, d'où le sentiment que le scamto est une langue crue qui exprime un certain manque de respect. D'un autre côté certaines appellations révèlent certains des problèmes sociaux que connaissent les townships.

3.7. *Métaphores lexicales*

La métaphore joue un grand rôle dans la manipulation lexicale en scamto. Le lexique des voitures en scamto est l'un des domaines riches en métaphores. Par exemple, la BMW série 3 E46 est couramment appelée *g-string* en scamto car, selon les locuteurs-créateurs, l'arrière de la voiture ressemble fortement au postérieur d'une femme portant un string. Le mot « *g-string* » (un sous-vêtement en anglais) devient alors une voiture en scamto. Un autre exemple est celui d'une Opel Kadett de 1994 à 2000 qui est appelée *ligundwane* en scamto à cause de

sa forme qui ressemble à un rat, le sens du mot en zoulou. Ces métaphores ont un côté amusant qui montre des interprétations plus humaines et sociales des machines et une certaine familiarité avec tout ce qui est moderne et nouveau. D'autres exemples qui ne sont pas du domaine de l'automobile existent aussi : *coconut* en scamto n'est pas la noix de coco comme c'est le cas en anglais mais une personne de couleur noire qui semble vouloir paraître blanche. D'autres métaphores concernent les noms des certaines personnes (souvent assez célèbres). On pourrait aussi appeler cette stratégie « antonomase » suivant la rhétorique ou « nom propre métaphorique ». Souvent les noms propres ne sont pas considérés comme des mots de la langue car ils n'ont pas de sens mais uniquement un référent qui est particulier. En scamto, les noms propres de certaines personnes deviennent des noms communs. Par exemple *Zola Budd*, le nom d'une ancienne marathonnienne et *Doctor Khumalo*, le nom d'un footballeur sont en scamto, des noms de voitures.

3.8. *Archaïsmes*

Nous retrouvons en scamto un retour aux clicks comme dans le nom *scamto*, ce parler qui fait objet de notre étude et qui comme nous l'avons souligné au début a d'autres appellations dans d'autres codes. Le « c » dans le nom *scamto* est un click dental, comme dans le mot *mca* (super). Les clicks qui existent dans les langues bantoues comme le xhosa, le zoulou et le sotho ont été hérités des langues khoïsan du pays, c'est-à-dire des langues des indigènes du pays. Ces langues qui étaient des langues à clicks ont en grande partie disparu dans le pays notamment à cause du changement vers l'afrikaans et d'autres langues bantoues. Ce changement a diverses motivations dont l'attitude des colons blancs qui ont jugé que ces langues étaient étranges et inarticulables. L'insertion de ce genre de sons dans une nouvelle langue pourrait être vue comme une expression de fierté identitaire de la part des locuteurs-créateurs et comme une sorte de retour aux sources. Les clicks sont en zoulou souvent dans des mots provenant des registres assez soutenus et de respect et qui ne sont pas utilisés dans des milieux urbains autant que dans des milieux ruraux. D'un autre côté, l'utilisation des clicks peut s'interpréter comme une forme d'exclusion – il est connu que ces sons sont difficiles à prononcer par les locuteurs d'autres langues comme les langues germaniques, ce qui fait donc du scamto une langue de la population noire.

3.9. *Emprunts*

La question des emprunts semble compliquée quand il s'agit de décrire un parler comme le scamto. Nous distinguons donc les mots héréditaires, qui sont ceux que le scamto a hérités des langues principales dans les townships, et les emprunts qui sont les mots qui viennent des langues très peu connues ou pas du tout dans les townships. Les emprunts en scamto gardent leur forme, par exemple, le mot *ola/hola/wola* qui vient de l'espagnol dont le sens ne change pas en scamto (salut !). D'autres emprunts subissent de petites modifications pour s'adapter au système du scamto, par exemple, *tcheri* qui vient du mot 'chéri' en français mais qui en scamto est utilisé exclusivement pour une petite-amie.

3.10. *Autres stratégies*

Il existe d'autres stratégies qui sont importantes dans l'innovation lexicale du scamto sans pour autant révéler beaucoup d'éléments identitaires.

3.10.1. Extensions sémantiques

Par exemple, le mot *mapapa* en scamto vient du sotho *papa* (un plat typique à base de la farine de maïs). En scamto il est employé à la forme plurielle et désigne la nourriture en général.

3.10.2. Onomatopée

Par exemple, *gusheshe*, le nom d'une voiture venant de l'onomatopée, « shesha » pour la chaleur.

3.10.3. Réduplications

Certains mots en scamto sont redoublés, par exemple, *toi-toi* (grève), *baza baza* (très grand/gros), *shap shap* (très bien). Les deux premiers ne peuvent exister que dans leur forme redoublée, tandis que pour le dernier le mot *shap* existe souvent seul.

Conclusion

Nous avons vu que la manipulation lexicale joue un rôle très important dans le développement du scamto. Le fait que le lexique du scamto est parallèle à celui ou à ceux du répertoire des locuteurs est la conséquence d'une manipulation lexicale et résulte d'un besoin d'exprimer une identité urbaine différente de celles qui existent déjà. Ces dernières sont souvent rurales et ne tiennent compte ni de la réalité des townships ni de la nouvelle réalité du pays. De nombreuses stratégies qui justifient la volonté des locuteurs dans la formation du lexique du parler sont mises en œuvre. Celles-ci révèlent des éléments de la nouvelle identité véhiculée par le scamto, notamment les manipulations de la forme comme l'hybridation, la relexification, la troncation, mais aussi des stratégies telles que les métaphores qui montrent non seulement un esprit créatif mais surtout une attitude positive envers tout ce qui est moderne et nouveau, sans pour autant mépriser les langues matrices qui sont les langues bantoues. Les stratégies les plus importantes sont donc la relexification et l'hybridation, deux principes fondamentaux dans la construction de codes secrets selon Bakker (1998), ce qui révèle le début du scamto : l'émergence d'un code secret dans certaines prisons sud-africaines, d'où le nom *tsotsitaal* langue (*taal* en afrikaans) de délinquants (*tsotsi* en sotho). Comme le déclare Golovko (2003 : 197), la manipulation lexicale ne suffit pas pour l'émergence d'une nouvelle langue car seulement certains mécanismes peuvent être contrôlés par les locuteurs. Mais si d'autres semblent être difficiles à contrôler, ils méritent pourtant tout autant d'être examinés aussi.

Références bibliographiques

- BAKKER, Peter (1997). "A Language of Our Own": The genesis of Michif, the mixed Cree-French language of the Canadian Métis. New York: Oxford University Press.
- BAKKER, Peter (1998). Para-Romani languages versus secret languages : Differences in origin, structure and use, in *The Romani Element in Non-Standard Speech*, Matras Yaron (Ed). Wiesbaden : Harrassowitz Verlag, 69-96.
- GOLOVKO, Evgeniy (2003). Language contact and group identity : The role of "folk linguistic engineering", in *The Mixed Language Debate : Theoretical and Empirical Advances*, Matras, Yaron & Bakker, Peter (Eds), Berlin, New York : Mouton de Gruyter, 177-207.
- LAFAGE, Suzanne (1998). Hybridation et 'français des rues' à Abidjan, in *Alternances codiques et français parlé en Afrique*, Queffélec, Ambroise (Ed). Université de Provence, 279-291.
- LEPAGE, Robert Brock, TABOURET-KELLER, Andrée (1985). *Acts of Identity, A creole based study of language and ethnicity*, Cambridge : Cambridge University Press.
- MOUS, Maarten (2003). The linguistic properties of lexical manipulation and its relevance for Ma'a in *The Mixed Language Debate : Theoretical and Empirical Advances*, Matras, Yaron & Bakker, Peter (Eds), Berlin, New York : Walter de Gruyter, 209-235.
- MUYSKEN, Pieter (1997). *Media Lengua*, in *Contact Languages, a wider perspective*, Thomason Sarah (Ed). Amsterdam/Philadelphia : John Benjamins, 365-426.

NICOLAI, Robert (2003). Contact et genèse : ouverture et perspectives: pour un “Nouveau Programme” de recherche sur l’évolution des langues, in XVIIth International Congress of Linguists, Prague.

NICOLAI, Robert (2005). Language processes, theory and description of language change, and building on the past : lessons from Songay, in *Linguistic Diversity and Language Theories*. Frajzyngier, Z, Hodges, A & Rood, D. S. (Eds). Amsterdam/Philadelphia : John Benjamins, 81-104.

NTSHANGASE, Dumisane (2002). Language and language practices in Soweto, in *Language in South Africa*, Mesthrie, Rajend (Ed). Cambridge : Cambridge University Press, 407-415.

DES REPRÉSENTATIONS DIGLOSSIQUES AUX PRATIQUES INTERLECTALES : CONFRONTATION DE CONCEPTS SOCIOLINGUISTIQUES EN CONTEXTE RÉUNIONNAIS

Séverine RAPANOËL

severine.rapanoel@wanadoo.fr

Lidilem, Université Grenoble 3

BP 25 – 38040 Grenoble

Abstract : This article examines the consequences of the appropriation of sociolinguistic concepts by teachers participating in a pilot educational program. It will shed light on the problems caused by the disparity between, on the one hand the representation of the sociolinguistic situation of Reunion Island based on the model of diglossia and, on the other hand, the real practices of speakers which reflect an interlectal mode of communication. These problems will be discussed and illustrated through the empirical data selected from our corpora.

L'ambition de cette contribution est de montrer les effets de l'appropriation de certains concepts sociolinguistiques notamment ceux de diglossie et d'interlangue par un groupe d'enseignants investis dans une action glottopolitique sur l'île de La Réunion centrée sur l'Apprentissage du Français en Milieu Créolophone (désormais l'AFMC).

L'AFMC est une action appuyée par un plan de formation des enseignants, menée sous la tutelle de l'Education Nationale en collaboration avec le projet éducatif global de la ville du Port. L'inspecteur, la conseillère pédagogique et la formatrice de ce projet organisent des stages pour des enseignants afin de les sensibiliser au contexte linguistique réunionnais et de les accompagner dans un changement de leurs pratiques langagières et pédagogiques en classe.

Le concept de diglossie s'est largement diffusé auprès des enseignants qui utilisent encore ce modèle pour saisir la réalité linguistique de La Réunion, alors que les faits langagiers analysés devraient aujourd'hui conduire à l'abandonner. L'appropriation de ce concept par les enseignants a cristallisé une représentation non valide des usages langagiers qui s'est répercutée dans les pratiques de classes et plus précisément dans les usages du créole dans ce contexte.

Dans un premier temps, il s'agit de revenir sur les concepts ayant permis de rendre compte des usages langagiers réunionnais – de la diglossie à l'interlecte – de critiquer ces concepts et d'établir pourquoi le concept de diglossie – mis en question en linguistique – reste celui auquel la sociolinguistique réunionnaise et l'AFMC se réfèrent.

Dans un deuxième temps, nous développerons la perception de la situation linguistique chez le groupe d'enseignants impliqués dans le programme AFMC en nous appuyant sur des entretiens semi-directifs réalisés entre 2002 et 2004. Puis, nous montrerons à travers des observations pédagogiques que les représentations des enseignants induisent des activités particulières basées sur une approche contrastive des langues française et créole et nient les pratiques langagières qu'ils utilisent habituellement en classe.

Nous rappelons que cette action glottopolitique se déroule dans un département français où le français remplit la fonction de langue officielle et de langue de l'école. Il occupe des positions symboliquement importantes et connaît une survalorisation dans les discours encore imprégnés de l'idéologie diglossique. Le créole, lui, n'est reconnu comme langue régionale que depuis 2001. Depuis, les discours à son égard sont plus modérés

qu'auparavant. Mais malgré cette politique de valorisation affichée, il reste toujours une langue minorée car il est continuellement comparé au standard dominant auquel on prête des fonctions socialement plus nobles.

1. L'avancée des travaux scientifiques : de la diglossie à l'interlecte

Dès les années 1970, les premiers travaux réalisés sur le terrain réunionnais font référence au concept de diglossie, élaboré par Charles Ferguson en 1959, pour traduire les usages langagiers des locuteurs. Ce modèle issu de la sociolinguistique nord-américaine rappelle qu'il existe deux langues qui se partagent l'espace énonciatif de façon fonctionnelle. Ainsi, le français est défini comme une langue prestigieuse, standardisée, apprise et non acquise (à l'exception de certaines familles, ce qui constitue déjà une entorse aux critères fergusoniens), utilisée dans des situations formelles. À l'opposé, le créole n'est pas considéré comme une langue, il est dévalorisé, non standardisé, acquis au sein de la famille, utilisé dans des situations informelles et on lui attribue des valeurs identitaires et affectives (c'est la langue du cœur). Même si depuis Charles Ferguson, le concept de diglossie a été critiqué et sans cesse revisité puisque sa définition jugée trop rigide camoufle les pratiques empiriques, il reste une référence majeure de la sociolinguistique réunionnaise car il souligne l'inégalité statutaire entre le français et le créole. Nombreuses sont les études qui dénoncent cette situation de diglossie et les conséquences qui en découlent telles que l'insécurité linguistique en français de l'enfant créolophone, les troubles psychiques comme le manque de confiance en soi, le sentiment d'infériorisation ou encore l'échec scolaire. Pierre Cellier, en 1985, dans sa recherche sur la description syntaxique du créole réunionnais, va même jusqu'à employer le terme de « dys- glossie ». Un terme qu'il utilise aussi, la même année, dans un article intitulé « dysglossie réunionnaise » paru dans la revue *Cahiers de praxématique*. Ce néologisme a pour intention de rappeler le dysfonctionnement sociolinguistique qu'impose la hiérarchie entre le français et le créole, entraînant par la même un conflit symbolique entre ces deux langues. Si le modèle diglossique connaît encore un certain succès aujourd'hui, c'est parce qu'il rappelle systématiquement la situation de domination d'une langue sur l'autre.

Cependant, d'autres approches ont été retenues pour rendre compte des usages langagiers à La Réunion et notamment celle du continuum linguistique. Le continuum complexifie la problématique du contact des langues qui, jusqu'ici, était traitée en termes de domination d'une variété haute sur une variété basse. Il postule l'existence d'une zone intermédiaire assimilable à une variation continue de mésolectes hiérarchisés répondant à des règles implicationnelles. Un lecte est caractérisé en quelque sorte par un faisceau de variantes liées par des relations d'implications c'est-à-dire que la présence d'une variante suppose nécessairement celle d'une autre. Les tenants de cette théorie ont pour objectif de démontrer que la variation n'est pas « libre » ou « accidentelle », mais régie par des règles.

Michel Carayol et Robert Chaudenson sont les premiers à avoir appliqué la théorie du continuum linguistique au contexte réunionnais dans les années 1970. En étudiant des variables phonétiques, ils montrent que l'ensemble des productions est situé sur un axe où la variété standard du français constitue le pôle acrolectal et, à l'autre extrémité, la variété du créole la plus archaïque et éloignée du français représente le pôle basilectal. Entre ces deux pôles, se trouve le mésolecte qui comporte deux variétés principales : le français régional et le français créolisé.

Il est vrai que l'usage du continuum présente des avantages au niveau de la description linguistique, mais il connaît aussi des faiblesses dues notamment à la rigidité des règles implicationnelles qui empêche la classification d'une quantité de productions langagières observées sur le terrain. Ainsi, il existe des usages langagiers qui ne trouvent pas leur place

sur l'axe du continuum comme les énoncés « métissés » où le créole et le français s'entremêlent ou encore des productions qui s'intègrent dans le moule syntaxique du français mais qui relèvent d'une prosodie caractéristique du créole. C'est pourquoi les derniers travaux sur la dynamique des langues en contact à La Réunion font de plus en plus appel aux concepts d'interlecte et de macro-système mis en évidence par Lambert-Félix Prudent, dès 1981, lorsqu'il a analysé la situation sociolinguistique martiniquaise.

Le macro-système constitué de l'ensemble des répertoires langagiers des Réunionnais offre la possibilité aux locuteurs de former des énoncés qui n'obéissent « ni au basilecte nucléaire ni à la grammaire acrolectale » (Prudent, 1981 : 26) et qui permettent toutes sortes d'« hybridations » ou de « mélanges » dans une infinie diversité. L'interlecte représente alors ce mode de communication où les Réunionnais ne cessent de mêler allégrement le français et le créole pour construire leurs discours en ne respectant pas les règles des deux langues puisque leurs paroles se forgent « sur un viol permanent des contraintes grammaticales des deux codes » (Prudent & Merida, 1984 : 43). Le nouveau regard que Lambert-Félix Prudent jette sur la situation sociolinguistique de son pays entraîne des bouleversements épistémologiques importants. Il récuse la notion de frontière entre le créole et le français. Il rejette donc la vision fergusonienne qui expose les langues comme des systèmes homogènes qui ne se mélangent pas et présente la « zone intermédiaire » comme « pouvant être autre chose que la manifestation fautive et provisoire du règlement d'un conflit de substitution » (Romani, 2000 : 31).

Ce sont toutefois les premiers travaux scientifiques qui continuent à trouver écho dans le secteur éducatif.

2. L'appropriation des concepts de diglossie et d'interlangue et les conséquences de cette appropriation sur les pratiques pédagogiques d'enseignants

Le concept de diglossie se répand dans les médias, dans les formations universitaires et professionnelles et, si les enseignants remettent aussi en cause la répartition fonctionnelle des usages face aux différentes situations de communication, ils s'appuient sur le modèle diglossique pour expliquer la domination du français sur le créole. Ils se servent de ce cadre de pensée pour lutter contre la minoration du créole à l'école et envisager une situation idéale, celle du bilinguisme.

Des enseignants participant au projet glottopolitique AFMC, axé sur l'apprentissage du français en milieu créolophone, se sont appropriés le concept de diglossie et y font régulièrement référence lors des entretiens que nous avons menés dans le cadre de notre thèse. Par exemple, lorsque nous avons interrogé une enseignante sur sa perception des langues présentes à La Réunion, elle a répondu :

E1 : D'un point de vue professionnel euh / ben il me semble que / il y a deux langues en présence hein / donc qu'on doit absolument tenir compte de l'une et de l'autre sans perdre de vue quand même que l'école doit être le >+ enfin le lieu où l'enfant va rentrer dans l'apprentissage de la langue nationale de la langue euh / de la réussite le français hein / donc bon maintenant le fait qu'il y ait le créole doit absolument être pris en compte ne serait-ce que pour que l'enfant ne soit pas en situation de diglossie quoi.¹

Dans la citation suivante, l'enseignante explique que l'exclusion du vernaculaire de l'enfant risque de maintenir la situation de diglossie :

E1 : Si on lui dit non c'est pas ta langue c'est pas celle-là qu'on parle dans la classe ben du coup il est dévalorisé par rapport à ce qui est dit dans sa famille / hein / c'est de la diglossie tout le monde en parle tout le monde connaît.

¹ Les conventions de transcription sont mentionnées à la fin de l'article.

Dans les entretiens que nous avons réalisés, nous avons pu constater que le point de vue de cette enseignante est largement partagé par ses collègues qui ont suivi la formation AFMC. Le concept de diglossie leur apparaît satisfaisant pour comprendre la situation sociolinguistique réunionnaise car ils n'en retiennent que deux traits principaux : l'existence de deux variétés homogènes et la hiérarchie entre ces deux variétés. Imprégnés profondément par une « idéologie monolingue », les enseignants ont du mal à saisir le dynamisme qui régit le système langagier réunionnais, leur conception d'une langue renvoyant à la définition traditionnelle commune d'une entité linguistique homogène. Par conséquent, par analogie avec le monolinguisme, leur représentation du bilinguisme se réduit à la capacité de parler deux langues dans des situations de communication précises sans les mélanger. D'un autre côté, les formes métissées que les enseignants observent dans les pratiques langagières ordinaires des élèves sont stigmatisées et considérées comme des sources d'erreurs, des interférences ou de l'« interlangue »².

E2 : Je crois qu'il y a un problème aujourd'hui c'est qu'on sait plus dans quoi on parle / c'est le problème de l'interlangue / et qu'on essaie nous de dépatouiller en classe / et donc ben ne sachant plus en quoi ils parlent ils parlent en tout / et ni le créole ni le français

L'analyse des enseignants sur les pratiques langagières des élèves postule l'existence d'un déficit linguistique. En effet, les enseignants identifient leurs élèves à des locuteurs « semilingues » ne sachant parler ni le français ni le créole car ils les mélangent et n'atteignent le niveau du « locuteur natif » dans aucune d'elles. Pour parvenir à développer le bilinguisme de leurs élèves, les enseignants appliquent une pédagogie basée sur la distinction des codes. Ils adoptent alors des attitudes et des pratiques professionnelles monolingues normatives qui risquent de ne pas produire les effets attendus. En effet, comment envisager une séparation nette entre le français et le créole dans un système langagier qui mêle ces variétés au point de rendre floues leurs frontières ? Les enseignants ne disposent pas de solution miracle puisqu'ils inscrivent eux-mêmes leurs pratiques dans une variation qui oscille entre le français et le créole. Ils affirment être aussi en « interlangue » :

1.EN : Et quand ils mélangent les deux qu'est-ce que vous faites ?

2.E3 : Alors là ça non / ça pas question / ou tu parles créole / ou tu parles français mais euh **mi** (*je*) euh >+ je **nana** (*ai*) / ça non / on ne veut pas / mais ça on en a pris conscience y a pas trop longtemps non plus hein / c'est toujours avec le stage AFMC / AFMC c'est l'apprentissage du français en milieu créolophone / donc elle nous a dit / ça vous n'acceptez pas >+ qu'ils s'expriment en interlangue / et qu'il faut qu'ils fassent bien la différence entre le français et le créole

3.EN : C'est eux qui vous ont dit de le faire en fait ? // donc tu leur dis quoi toi ? / de parler en français ? < +

4.E3 : De dire en français / ou de dire en créole mais de ne pas mélanger les deux / et ça on essaie de sensibiliser les parents aussi à ça parce que / si à la maison c'est pareil tu auras toujours euh // >+ ce qui nous mène là >+ d'ailleurs nous on le fait / les enseignants mélangent allégrement français / créole / on est souvent dans l'interlangue // parce qu'on ne parle pas le même créole que les enfants déjà / on ne parle pas ce créole basilectal là / bon / et puis comme

² Nous avons mis des guillemets à ce terme afin de montrer que les enseignants, conscients des spécificités communicatives réunionnaises où les locuteurs mélangent les langues, ont tendance à amalgamer ce que les chercheurs définissent comme l'interlecte et l'interlangue. Il est hors de doute que les productions issues du contact de langues partagent certains traits avec ce qu'il est convenu d'appeler l'interlangue des sujets en situation d'apprentissage. L'interlecte et l'interlangue cependant ne relèvent pas de la même logique communicationnelle. Dans de nombreux cas, distinguer l'interlangue de l'interlecte n'est pas chose aisée même si nous pouvons tout de même mentionner quelques différences :

- l'interlangue se situe dans un processus d'apprentissage tandis que l'interlecte est une stratégie de communication ;

- l'interlangue est une étape provisoire vouée à disparaître dans le temps (lorsque le sujet maîtrisera mieux la langue cible) tandis que l'interlecte s'inscrit dans la durée (elle persistera tant que les locuteurs éprouveront du plaisir à jongler avec les langues).

on a plus le français tu vois dans notre >+ donc on a un petit peu de mal / à parler QUE le créole PUR ou jamais // on a les deux cultures / donc on mélange les deux

5.EN : Et par contre vous ne l'acceptez pas chez l'enfant ?

6.E3 : Ben chez l'enfant / non // non

Dans cet extrait, l'enseignante montre bien la difficulté qui se crée lorsque l'objectif de langage (la séparation des langues) est confronté aux pratiques empiriques (mélange des langues). On peut alors se demander comment elle parviendra à atteindre ses objectifs de langage si elle-même ne les applique pas. L'enseignante ne s'assimile pas aux apprenants que sont ses élèves et dès lors le fait qu'elle mélange le français et le créole ne se révèle pas problématique pour elle. La compétence qu'elle a acquise dans les deux langues lui permet de les mélanger, ce qui n'est pas le cas des élèves, lesquels peuvent subir en outre l'influence du milieu familial. Toutes ces confusions montrent qu'il est difficile de raisonner et d'agir dans le cadre archétypique de la diglossie pour expliquer et réglementer les usages langagiers à La Réunion. Par conséquent, il paraît souhaitable dans une perspective de changement des pratiques et donc de formation de sensibiliser les enseignants à la variation linguistique et de leur démontrer que l'interlecte est un mode de communication répandu chez les locuteurs réunionnais. Il ne doit donc pas être interprété comme des aberrations fautives produites par des personnes ne sachant parler ni le français ni le créole, mais comme une volonté de communiquer avec toutes les formes disponibles de leur répertoire langagier comprenant le français dans toute ses variations, le créole dans toute ses variations et tout le panel linguistique situé entre ces deux pôles.

Actuellement, les enseignants engagés dans le programme AFMC sont marqués par une conception sociolinguistique fondée sur la diglossie et se basent sur ce modèle pour tenter de développer une compétence bilingue idéale chez leurs élèves. Ils n'accordent aucune place aux pratiques effectives des locuteurs réunionnais. Ils calquent leur pédagogie du créole sur celle du français et tombent dans les mêmes dérives, à savoir : ils privilégient la forme et délaissent la composante communicative en ne proposant pratiquement que des activités contrastives. Ils exigent que les élèves reproduisent des énoncés créoles de la variété basilectale et rejettent les autres variétés de cette langue.

Nous proposons maintenant d'analyser un extrait de notre corpus d'observations de classes³ qui illustre les pratiques types des enseignants qui ont voulu introduire le créole dans les apprentissages suite à la formation AFMC.

1.EL1 : Je parle français

2.M : Tu parles français / qu'est-ce que tu veux ?

On s'aperçoit en 1 que le but de la séance est la distinction des langues qui doit être conscientisée dès la prise de parole par la verbalisation du choix de la langue dans laquelle on va communiquer. Au cours de l'activité, l'élève a le choix entre s'exprimer en français ou en créole et l'enseignante refuse les « mélanges ». Au tour de parole 6, l'enseignante respecte les exigences de la séance car elle ne valide pas l'énoncé de EL1 qui a choisi de s'exprimer en français et qui a employé « gagner » au lieu de « avoir ».

3.EL1 : Je veux l'enfant qui a gagné des ballons

4.M : L'enfant qui ? // ah vous avez entendu ce qu'il a dit ?

5.GEL : Oui

6.M : Il a mélangé français créole / alors tu as dit que tu parlais français / alors ? // l'enfant qui A le ballon / allez

³ Notre corpus d'observations de classes des enseignants ayant pratiqué l'AFMC comporte 106 heures d'enregistrement dans cinq classes différentes des trois cycles d'enseignement.

Par contre, au tour de parole 9, EL2 produit un énoncé créole correct mais l'enseignante ne le valide pas car EL2 utilise le terme « zanfán », expression bien créole mais à consonance plus française que « marmay ».

9.EL2 : **Mi ve le zanfán la giny son >+ son chosèt**

(Je veux l'enfant qui a eu ses >+ ses chaussettes) [...]

13.M : Alors comment on va dire ?

14.EL3 : **Le marmay**

(L'enfant) [...]

16.M : **Mi ve le marmay la giny chosèt** // ah beh t'as perdu au fait / on va retourner tout à l'heure

(Je veux l'enfant qui a eu des chaussettes // ah beh t'as perdu au fait / on va retourner tout à l'heure)

Au tour de parole 25, l'enseignante ne valide pas la forme verbale « la brulé ». En créole, il existe deux formes verbales, la forme longue « la brulé » relevant de la variété acrolectale et la forme courte « la bril » qui appartient à la variété basilectale. Donc, les deux productions sont correctes, mais dans les représentations de cette enseignante, la variété basilectale semble la plus appropriée pour représenter la norme du créole. C'est pourquoi, elle rejette les autres formes.

24.EL4 : **Mi ve le marmay la brulé son min**

(Je veux l'enfant qui a brûlé sa main)

25.M : Redis lui Nicolas

26.EL3 : **Mi ve le marmay la bril son min**

(Je veux l'enfant qui a brûlé sa main)

27.EL4 : **Le marmay la BRIL son min**

(L'enfant qui a BRULE sa main)

Au tour de parole 46 et 65, la maîtresse, qui doit servir de modèle pour cette séance basée sur la distinction des langues, transgresse elle-même les règles puisqu'elle passe d'une langue à l'autre.

46.M : Mathieu la pris / oh la la / la **mayé** là hein / alors qu'est-ce qu'il te fallait toi ?

(Mathieu a pris / oh la la / on s'est trompé là hein / alors qu'est-ce qu'il te fallait toi ?)

65.M : Attends / t'avais dit que tu **koz** créole / **mi ve le marmay**

(Attends / t'avais dit que tu parles créole / je veux l'enfant)

Enfin, au tour de parole 48, l'enseignante identifie mal l'énoncé car elle dit à l'élève qu'il a parlé français alors qu'il s'est exprimé dans les deux langues.

47.EL3 : **Mi koz kréol / mi ve** l'enfant qui est en colère

(Je parle créole / je veux l'enfant qui est en colère)

48.M : Non / tu as parlé français là

Toutes ces pratiques sont bien révélatrices d'un tâtonnement empirique où l'enseignante s'efforce de raisonner selon un schéma linguistique très simplifié, caractérisé par un face à face entre deux langues homogènes bien déterminées. Son but dans les séances AFMC est de placer les langues au même niveau. Or, elle semble oublier que le créole n'a pas les mêmes règles que le français. C'est une langue à tradition orale et ce paramètre devrait être pris en compte dans l'optique de son intégration dans les apprentissages. À notre avis, il n'est pas logique de vouloir transposer la pédagogie d'une langue standardisée et normée pour l'appliquer méthodiquement à une langue qui ne dispose pas encore de standard. De plus, il nous paraît inefficace de se fonder sur le modèle « caricatural » de la diglossie pour vouloir le dépasser et arriver à la situation désirée de bilinguisme.

Pour conclure, nous observons que le programme AFMC a sensibilisé les enseignants à la nécessité de prendre en compte le créole dans les apprentissages et les a encouragés à initier des activités en classe. Cette langue minorée trouve ainsi une place revalorisée à côté du

français. Cependant, par manque d'encadrement, de moyens et par la transmission d'un savoir théorique non valide, les enseignants se sont lancés dans une pédagogie désorganisée. Ils tâtonnent pour trouver des solutions afin de développer un bilinguisme parfait (qui n'existe pas) chez leurs élèves. Ils mettent en place des activités contrastives en pensant que la distinction des deux langues permettra aux élèves de mieux développer leur conscience linguistique. Ils privilégient alors la variété basilectale la plus éloignée du français pour bien marquer l'éloignement des deux langues. On peut se demander si l'adoption de tels comportements ne va pas induire, à la longue, une insécurité linguistique des élèves en créole et, le cas échéant, quel peut être l'intérêt d'introduire le créole à l'école.

Il est donc nécessaire de penser autrement la pédagogie du créole et de ne pas se contenter d'un transfert pédagogique calqué sur l'apprentissage du français qui n'a d'ailleurs pas fait ses preuves dans les contextes créolophones puisque l'échec scolaire est encore important. Ne serait-il pas temps maintenant dans les formations, de véhiculer une image plus dynamique des langues qui se côtoient à l'intérieur du répertoire langagier de chaque locuteur réunionnais, de montrer les usages diversifiés de leurs pratiques et d'entreprendre l'apprentissage des langues et de leurs normes en tenant compte du mode spécifique de communication lié à l'interlecte ?

Conventions de transcription

/	Pause brève, correspondant à la fin de la chaîne parlée
//	Pause moyenne ou longue (la longueur n'est pas objectivée par une mesure de temps)
>+	Auto-interruption ; rupture de construction, faux départ
<+	Hétéro-interruption : le locuteur est interrompu par une intervention d'un tiers
chanTÉ	Phonème, syllabe ou segment accentué
X	Syllabe inaudible
XX	Deux ou plusieurs syllabes inaudibles
(rires)	Phénomènes non verbaux, commentaires du transcripteur
?	Interrogation marquée ou non marquée morphologiquement
EN	Enquêtrice
E	Enquêté
M	Maître / Maîtresse
EL	Elève
GEL	Groupe d'élèves

Références bibliographiques

- BAVOUX, Claudine (2002). Représentations et attitudes dans les aires créolophones, in *Linguistique et créolistique*, Bavoix Claudine & De Robillard Didier (Dir). Univers créoles 2, Paris : Anthropos, 57-76.
- CARAYOL, Michel & CHAUDENSON, Robert (1978). Diglossie et continuum linguistique à La Réunion, in *Les Français devant la norme*, Gueunier Nicole *et al.* Paris : Champion, 175-200.
- CARAYOL, Michel & CHAUDENSON, Robert (1979). Essai d'analyse implicationnelle d'un continuum linguistique in *Plurilinguisme, normes situations, stratégies*, Manessy Gabriel & Wald Paul (Eds). Paris : L'harmattan, 129-174.
- CELLIER, Pierre (1985a). *Comparaison syntaxique du créole réunionnais et du français*. Saint-Denis : Université de La Réunion.
- CELLIER, Pierre (1985b). Dysglossie réunionnaise, *Cahiers de praxématique*, 5 : 45-64.
- FERGUSON, Charles A. (1959). Diglossia, *Word*, XV : 325-340.
- PRUDENT, Lambert-Félix (1981). Diglossie et interlecte, *Langages*, 61 : 13-38.
- PRUDENT, Lambert-Félix & MÉRIDA, Georges-Jacques (1984). An langaj kréyol dimipanaché...interlecte et dynamique conversationnelle, *Langages*, 74 : 31-45.

ROMANI, Jean-Paul (2000). L'interlecte martiniquais : Approches sociolinguistiques des rapports langue-idéologie dans une communauté antillaise. Thèse de Doctorat en Sciences du Langage. Université de Rouen Haute Normandie.

COMPÉTENCE PLURILINGUE, REPRÉSENTATIONS DES LANGUES ET CONSTRUCTION DES CONNAISSANCES CHEZ LES APPRENANTS FRANCO-ALLEMANDS

Sofia STRATILAKI

sofiastratilaki@web.de

Université Paris 8-Vincennes et DILTEC (Paris 3)

2, rue de la Liberté

93526 Saint-Denis Cedex

Abstract : References to the concept of representation are becoming increasingly common in studies on language learning and use. Of particular concern for the present paper are learners' representations of plurilingual competence and their motivation and attitudes towards languages. We argue that representations are developed in a complex interaction involving different languages and learners.

Introduction

L'intérêt relativement récent porté par la didactique des langues au plurilinguisme s'est accompagné d'une prise en considération attentive des phénomènes d'alternance des langues, tant dans l'usage ordinaire des locuteurs bi-/plurilingues que dans l'usage des locuteurs/apprenants eux-mêmes placés, du fait de leur biographie langagière ou de leur mode de scolarisation, dans des environnements bi-/plurilingues valorisants. Or, si l'alternance codique peut être reconnue comme un phénomène faisant partie de l'expérience communicative des participants dans les échanges exolingues hors contexte institutionnel, sa pratique à l'intérieur de l'école peut être jugée légitime ou illégitime. En ce sens, elle peut être envisagée soit comme symptôme d'incompétence et, dans ce dernier cas, elle est ressentie négativement comme la formulation des énoncés fautifs relevant d'interférences, soit elle est perçue comme un outil discursif facilitant la communication et faisant partie des stratégies d'exploitation du répertoire bi-/plurilingue par des sujets bilingues en cours d'apprentissage et fortement déséquilibrés dans leurs connaissances de l'une et l'autre langues.

Une telle conception de l'alternance des langues, envisagée sous l'angle bilingue, est particulièrement marquée à l'intérieur du système scolaire des lycées franco-allemands de Buc (Versailles) et de Sarrebruck (Sarre)¹, dont les principes pédagogiques se fondent sur la reconnaissance du changement des langues en tant qu'une ressource discursive complexe et sophistiquée aux fonctions multiples. Dans ces contextes éducatifs, la légitimité de la micro-alternance, qui suppose la mise en contact de deux langues au moins dans le répertoire langagier de l'élève franco-allemand, comme trait constitutif de sa compétence plurilingue et

¹ Le travail présenté ici se situe dans le cadre d'une recherche de plus ample envergure portant sur les représentations et les pratiques langagières plurilingues dans les Lycées franco-allemands de Buc (Versailles), de Fribourg (Brisgau) et de Sarrebruck (Sarre). Il importe de souligner que notre recherche se base sur un corpus large de données orales et écrites en allemand et en français recueillies par le biais des questionnaires et des entretiens semi-directifs conduits avec les élèves dans les trois lycées. Pour des raisons d'espace ne permettant pas de présenter ici la totalité de ce corpus, nous nous limiterons à l'analyse des données orales de deux élèves des lycées de Buc et de Sarrebruck. Pour les mêmes raisons, nous nous permettons de renvoyer à Stratilaki (2004a), si l'on souhaite étayer la réflexion sur l'organisation curriculaire de ces lycées.

comme stratégie ordinaire caractérisant le parler bilingue, semble être largement reconnue et exploitée, souvent comme moyen pour acquérir de nouvelles compétences et connaissances. La prise en compte des phénomènes d'alternance des langues dans l'appropriation et la construction des savoirs invite donc à se pencher, de manière plus précise, aussi bien sur les représentations des apprenants du bi-/plurilinguisme individuel et sociétal que sur leur répertoire langagier développé à travers des interactions sociales à l'école comme en milieu ordinaire.

À partir de ces réflexions, l'articulation que le présent article entend approfondir est celle entre, d'une part, des pratiques langagières observables dans les deux établissements et des représentations repérables du plurilinguisme et, d'autre part, des compétences plurilingues sous-tendant de telles pratiques et étant pour partie marquées par ces représentations. À cette fin, nous analyserons les apports réciproques des représentations du bi-/plurilinguisme et des contextes d'acquisition en termes de construction et de gestion de la compétence bi- ou plurilingue chez deux apprenants franco-allemands. Nous partirons plus particulièrement du constat que la construction de la compétence plurilingue chez les apprenants dans les lycées binationaux contient deux images bien distinctes, plus ou moins soutenues par des configurations de réseaux sociaux des locuteurs au sein desquels cette construction se déploie. Il s'agit, d'un côté, de l'image d'une compétence plurilingue enthousiaste, à orientation internationale, caractérisée par une grande mobilité, et par conséquent par une moindre stabilité. Elle est, de ce fait, susceptible d'évolutions s'organisant en fonction des combinaisons complexes, différentes pour chaque apprenant, et adaptées aux besoins de celui-ci pour chaque situation. À cet égard, le lycée franco-allemand de Buc se présente potentiellement comme une institution jouant un rôle clef dans l'évolution de la compétence plurilingue de l'élève. De l'autre, il s'agit d'une compétence bilingue consciente et fière de l'être, issue du milieu familial et frontalier et dès lors plus stable et solidaire. De ce point de vue, l'établissement binational de Sarrebruck joue un rôle déterminant dans la scolarisation des apprenants en termes de valorisation d'un bilinguisme régional où l'appropriation des langues devient un atout social, voire économique. Ainsi, alors que le voisinage frontalier et le contact avec des locuteurs natifs de la langue apprise aident à la construction d'une compétence bilingue chez les apprenants de Sarrebruck, l'articulation entre différentes formes de mobilité géographique et de mobilité langagière chez les apprenants de Buc contribue à la construction d'une relation suivie et intense avec plusieurs langues favorisant une ouverture au plurilinguisme, en particulier la reconnaissance du caractère déséquilibré et pluriel de la compétence plurilingue.

Aborder la dynamique de la compétence plurilingue, cela signifierait alors prendre en considération les représentations des apprenants, leurs pratiques langagières ou encore leur mobilité langagière et questionner leur rôle dans le processus d'appropriation des langues et des savoirs. C'est donc à ces réflexions que seront consacrés les propos qui suivent.

1. La notion de compétence plurilingue et ses liens avec les représentations du bi-/plurilinguisme

L'organisation de notre analyse, après formulation de nos options majeures, répond à l'ordre suivant. Dans un premier temps, nous rappellerons au lecteur certains travaux théoriques innovants dans le champ de recherche sur le plurilinguisme tout en soulignant l'importance de la notion de compétence plurilingue pour notre étude de cas. Nous préciserons ensuite certains aspects de cette notion, afin de mettre en évidence ses liens avec les représentations que les apprenants ont sur le plurilinguisme. Enfin, nous décrirons brièvement, à différents niveaux, les liens que la compétence plurilingue entretient avec les pratiques langagières de deux élèves franco-allemands.

1.1. Compétence plurilingue et représentations des langues : quels liens ?

La notion de compétence plurilingue, telle qu'elle a été élaborée par Coste, Moore & Zarate (1997) et retenue dans le *Cadre Européen Commun de Référence* (2001), est à considérer à la fois en termes de contacts de langues et de parcours d'expérience, mais aussi en tant que potentiel de gestion du répertoire plurilingue dans les aspects dynamiques, langagiers et identitaires qui en sont constitutifs. Il s'agit plus particulièrement d'une compétence qui est à mettre en relation avec un itinéraire d'apprentissage impliquant des modifications d'équilibre, et étant de nature hétérogène, évolutive et changeante au cours de la vie des apprenants (Coste 2004 ; Stratilaki 2004b, 2005). En effet, la désignation de compétence plurilingue porte sur la compétence discursive d'acteurs sociaux en mesure de communiquer, à des degrés variables dans des langues différentes. À ce titre, il s'agit d'une compétence dynamique, dans le sens où ces acteurs sont à même de gérer et remodeler cette compétence plurilingue au cours de leur trajectoire personnelle, en fonction de leurs besoins et de situations sociales et discursives diverses (Stratilaki 2006 ; Moore 2006). La compétence plurilingue est ainsi posée comme un ensemble structuré, plutôt que segmenté, de compétences partielles diversement acquises et sollicitées renvoyant à la maîtrise plus ou moins limitée que possède un acteur, à un moment donné, dans une langue ou une autre.

Ces considérations sommaires nous permettent de faire deux observations. La première renvoie au fait que la compétence plurilingue est caractérisée par un répertoire pluriel de composantes plus ou moins développées quant aux capacités communicationnelles qu'elles présentent. Dans le contexte de l'école, où les représentations fonctionnent comme des filtres d'apprentissage et orientent les pratiques de langues, c'est l'ensemble de ce répertoire qui est à prendre en compte et à valoriser dans la construction de la compétence plurilingue chez les apprenants. La deuxième tient au fait que la compétence plurilingue se présente, de manière générale, comme *plurielle* (du bi- au pluri-) et *déséquilibrée*. En raison de ce déséquilibre, la configuration de cette compétence évolue, s'enrichit de nouvelles composantes, en complète ou transforme d'autres. Cette hétérogénéité peut être valorisée par les institutions bilingues comme une richesse contribuant à complexifier la compétence bilingue déjà acquise par les locuteurs, soit dans leur milieu familial, soit dans une région frontalière.

Avant d'aller plus loin dans la discussion, quelques réflexions s'imposent par rapport aux stratégies discursives des apprenants plurilingues. La compétence plurilingue se caractérise par le fait que, dans sa mise en œuvre, les stratégies mobilisées par l'apprenant pour la réalisation des tâches à dimensions langagières peuvent varier suivant les langues ou les combinaisons de langues auxquelles l'apprenant en question a recours, des représentations qu'il se fait de ses ressources (représentations individuelles et collectives, par exemple), ou encore des représentations qu'il a sur les ressources langagières des (inter-)locuteurs. En d'autres termes, l'apprenant peut avoir recours à des formes d'alternance codique, lorsque la représentation qu'il a sur les ressources de son interlocuteur lui permet de passer d'une langue à l'autre dans une même séquence d'échange. Il importe par ailleurs de souligner que la compétence plurilingue autorise des combinaisons, des alternances ou des jeux de langues sur plusieurs niveaux. À ce titre, il est possible de procéder à des changements de codes au cours d'un échange et de recourir à des formes de parler bilingue orientées, soit sur la continuation de déroulement de la communication, soit sur les moyens linguistiques nécessaires à la transmission des informations. En ce sens, le répertoire plurilingue de l'apprenant, considéré comme un ensemble riche de variétés langagières et d'options à disposition, permet des choix multiples reposant sur la variation interlinguistique, lorsque les circonstances le permettent (Aymonod *et alii*. 2006).

En raison de cet ensemble complexe de propriétés, la compétence plurilingue englobe une autre notion, celle de *compétence partielle*, comme le rappellent Coste, Moore & Zarate :

« une compétence acquise dans une langue est partielle dans la mesure où elle est partie d'une compétence plurilingue qui l'englobe et dans la mesure où, s'agissant de cette langue, elle "encapacite" l'acteur considéré pour certaines activités langagières ou pour certains contextes d'usage plus que pour d'autres activités ou d'autres contextes d'usage ». (Coste, Moore & Zarate, 1997 : 14)²

La compétence partielle peut concerner des activités langagières, des tâches spécifiques dans un domaine particulier, mais elle peut également avoir trait à des compétences générales individuelles. Plus spécifiquement, l'apprenant plurilingue dispose des stratégies métalangagières, c'est-à-dire des stratégies lui permettant d'utiliser la diversité des savoir-apprendre qu'il a à sa disposition pour établir des liens, imaginaires ou réels, entre les langues de son répertoire. C'est à travers ces expériences diverses de gestion des langues, dès lors que celles-ci ne sont ni cloisonnées entre elles ni strictement répétitives, qu'il enrichit ses capacités à apprendre³. Face à la complexité de cette situation, l'apprenant met en œuvre des stratégies où la proximité et la distance, effectives mais surtout représentées, entre les langues vont jouer un rôle important pour acquérir des connaissances nouvelles. Dans cette perspective, l'appréciation de la distance et/ou de la proximité linguistiques devient un facteur décisif dans la construction et la gestion de la compétence plurilingue.

La construction de la compétence plurilingue, que nous avons essayé de caractériser brièvement, provient, à notre sens, soit d'une scolarisation dans un établissement bilingue, soit d'une expérience « ordinaire » de la pluralité des contacts de langues (et des langues en contact) des apprenants. Nous allons maintenant nous intéresser aux effets du contact des langues, tels qu'ils se matérialisent dans les pratiques langagières de deux élèves franco-allemands, en tant qu'acteurs sociaux et *locus* du contact.

2. Le rôle de l'alternance codique dans les pratiques langagières des apprenants franco-allemands

Notre premier témoin a 12 ans. Cette élève, de nationalité française, fréquente le lycée franco-allemand de Buc depuis 2 ans. Auparavant, elle a été scolarisée dans une école maternelle à Hambourg et dans une école primaire à Vienne. Elle a ensuite été scolarisée au lycée international de St Germain en Laye, en France. Elle passe souvent ses vacances en Turquie et va tous les ans en Allemagne pour des séjours linguistiques et sportifs. En regardant de plus près la biographie langagière de l'élève, nous remarquons qu'elle considère comme langue maternelle le français, puisque toute sa famille parle français. Cela ne l'empêche pas d'alterner les langues au cours de la communication. Elle parle surtout le français, et parfois l'allemand, avec son père. Elle a commencé à apprendre l'allemand en Allemagne avec ses parents, et notamment avec son père, qui est professeur d'allemand, mais elle préfère garder le français en tant que langue unique de communication avec sa mère, son frère et ses camarades à l'école. Elle apprend l'anglais au lycée de Buc et à la maison, avec l'aide de son père. Selon elle, l'anglais ressemble beaucoup au français, même si la prononciation diffère dans les deux langues. Nous constatons, à travers le témoignage de cette élève, que des rapports complexes sont tissés entre le français, l'anglais et l'allemand. L'élève alterne les langues au cours de sa communication, nous l'avons dit, et ce malgré la maîtrise diversifiée de différents systèmes linguistiques dans son répertoire langagier. Cette

² *Op.cit.* Sur la notion de compétence partielle, voir également Stratilaki (2007, à paraître).

³ Cf. Conseil de l'Europe (2001).

représentation est renforcée au sein du milieu familial, où l'apprentissage des langues est perçu comme valorisant, comme le témoigne l'extrait suivant :

Extrait A⁴

Sandrine : « *Je préfère communiquer dans toutes les langues allemand & français & anglais car je suis assez à l'aise (...) En anglais/ j'aime & car c'est amusant\ même si je me trompe tout le temps (...) y changent la prononciation du mot & même si c'est le même mot\ Les deux langues se ressemblent mais elles ne sont pas les mêmes\ Je change les langues de communication\ lorsque j'écris à ma correspondante\ j'écris une phrase en français puis une en allemand.. quelques mots en anglais\ Et même/ si je suis bonne il me manque encore du vocabulaire (...) J'ai appris le français dans la famille car toute ma famille parle français\ l'allemand c'était avec mon père car il a été prof d'allemand\ dans l'école/ j'apprends l'anglais et je le parle avec mon père.. un peu... avec ma mère aussi parfois\ Elle apprend aussi l'allemand...et on rigole pas mal* ».

Sa principale difficulté en allemand et en français est la grammaire. En anglais, elle rencontre des difficultés dans l'apprentissage du vocabulaire et la prononciation des mots composés. Enfin, il est intéressant de remarquer le manque de confiance chez l'élève à l'égard de sa compétence plurilingue, ce qui n'empêche pas une ouverture au plurilinguisme et aux alternances codiques, comme nous le constatons en lisant l'extrait suivant :

Extrait B

Sandrine : « *On a un choix plus important pour les métiers\ les études car parler deux langues c'est mieux (...) Inconvénients/ il N'y en a PAS...je n'ai pas choisi\ mais c'est seulement car mon père est venu travailler au LFA (...) j'aime l'allemand/ les langues\ L'école et la famille sont les meilleures situations pour apprendre des langues....bilinguisme pour MOI c'est de savoir parler les deux langues sans aucun problème/ pouvoir communiquer et s'entendre avec des personnes\ qui parlent l'allemand (...) Non je ne me considère pas comme bilingue\ car j'ai quand même un peu de mal en allemand\ même si je suis très bonne mais il me manque encore du vocabulaire* ».

Nous passons maintenant à l'analyse de la biographie langagière de la deuxième élève. Ursula a 17 ans, elle est de nationalité française et actuellement en première littéraire de la section française au lycée franco-allemand de Sarrebruck. Voici un extrait de son témoignage :

Extrait C

Ursula : « *À l'oral/ pour des choses banales/ en anglais\ & car j'aime bien les accents de cette langue\ Pour exprimer mes sentiments\ le français/ est une langue plus belle.. À l'écrit/ je préfère le français & car en allemand/ j'ai trop des problèmes\ dans le genre des mots\ et donc des déclinaisons\ Mon vocabulaire n'est pas assez riche/ par rapport au français.. En français/ j'ai une mauvaise connaissance des outils techniques\ pour l'analyse littéraire\ (...) J'aime beaucoup le français et je trouve la philosophie intéressante (...) Difficultés/ den schwierigen WORTSchatz⁵/...Les langues déjà ne doivent pas être uniquement apprises dans le cadre de l'école\ Leur apprentissage peut être largement favorisé par les amitiés\ ou lors des échanges avec des correspondants\ Le fait de se retrouver dans un pays où l'on parle une*

⁴ Les transcriptions sont orthographiques. Conventions principales de transcriptions adoptées : pauses petite/moyenne, l'' pauses en secondes, : allongement de la syllabe, /\ intonations montante/ et descendante, PLUringue segment accentué, xxx segment inaudible, & enchaînement rapide, (...) passage entre les informations, ^ liaison. Remarques complémentaires : les deux apprenants sont caractérisés par un pseudonyme.

⁵ Notre traduction en français : « [...] le vocabulaire est difficile ».

langue étrangère encourage la communication\ Enfin\ le mieux serait bien sûr d'avoir déjà des parents bilingues &ou de nationalité différente.. pour bénéficier d'une culture mixte\ Je préfère communiquer en français et en allemand, plus rarement (...) En général/ non/ je ne change pas la langue de communication\ pas vraiment\ &sauf s'il y a des problèmes de compréhension\ on passe dans l'autre langue\ Je parle en français\ même avec mes amis bilingues de la section allemande\ Il peut arriver que je parle en allemand &mais c'est plutôt rare\ Je ne fais pas beaucoup de séjours en Allemagne/ sauf aller à l'école bien entendu\ mais il y a déjà des échanges dans le cadre du lycée qui ont été assez bénéfiques\ ...Je n'ai pas fait des séjours d'une ou 2 semaines en Italie & Autriche & au Danemark & en Hollande/ mais j'appris à parler l'italien ».

Ursula parle le français avec ses parents et sa sœur, le français et l'allemand avec ses grands-parents. Comme langue maternelle, elle désigne le français. Néanmoins, les langues préférées de l'élève sont le français et l'anglais, selon des critères bien distincts : pour l'anglais, sa préférence est plutôt ancrée sur l'aspect langagier « *accent de cette langue* » et pour le français sur l'aspect affectif « *belle langue* ». Il paraît ainsi que sa compétence plurilingue se caractérise par un déséquilibre de maîtrise en fonction des langues (allemand/français) entraînant une préférence à l'écrit pour le français, et ce malgré les difficultés rencontrées. L'importance de la construction de la compétence plurilingue chez l'élève a été marquée par le choix de ses parents d'une scolarisation bilingue dans l'établissement franco-allemand. La mobilité géographique de l'élève reste surtout nationale. Nous remarquons que ses déplacements ont été motivés par des raisons aussi bien extrascolaires (vacances) que scolaires (séjours linguistiques). Sur le plan linguistique, il est à noter que sa difficulté principale dans l'apprentissage de l'allemand relève du domaine lexical. Ursula n'a pas souvent recours à des formes d'alternance codique. En effet, elle préfère parler en français même avec ses amis germanophones. C'est la raison pour laquelle elle alterne les langues seulement aux moments des lacunes lexicales ou même lorsqu'elle rencontre des problèmes durant la communication. Il importe de souligner que l'intérêt de cet extrait se trouve aussi dans la référence au biculturalisme, acquis dans une famille binationale et progressivement renforcé par la construction d'une compétence plurilingue chez notre témoin. Pour conclure, nous pouvons dire que l'élève montre un esprit ouvert à l'apprentissage des langues dans et hors de l'école « *amitiés, correspondance* ». Elle tire profit de ses séjours à l'étranger pour développer sa compétence plurilingue, même si ceux-ci sont essentiellement de courte durée.

Si nous nous prêtons brièvement au jeu d'un regard rétrospectif sur les analyses effectuées, nous constatons qu'à la différence du bilinguisme, qui mobilise majoritairement des images d'équilibre entre deux entités équivalentes et qui ne constitue pas une condition suffisante pour le développement des compétences linguistiques, le plurilinguisme, entendu comme au moins deux langues, implique en effet constitutivement le déséquilibre et favorise le développement des compétences diversifiées chez les élèves franco-allemands. Dans cette perspective, selon la trajectoire de l'apprenant, la configuration de cette compétence évolue, s'enrichit de nouvelles composantes, en complète ou transforme certaines autres, en laisse encore certaines autres déperir. Ces changements ne traduisent pas, à notre sens, un déficit au niveau des capacités ou des performances. Ils constituent au contraire une *ressource supplémentaire* dont disposent les élèves franco-allemands et dont ils se servent pour communiquer du sens et pour construire du savoir.

En guise de conclusion

À l'issue des analyses effectuées, quelques points méritent d'être soulignés à l'égard des représentations et des pratiques langagières des apprenants franco-allemands. Nous remarquons, en premier lieu, que les représentations fonctionnent comme des mécanismes

qui aident les élèves à gérer le choix des langues au cours de l'interaction et à interpréter leur sens et leur valeur conversationnelle. En ce sens, elles permettent aux élèves de procéder à des opérations mentales relativement complexes dans le cadre d'une analyse métalinguistique, tels que le repérage d'éléments communs entre les langues, l'alternance codique ou même le recours à des connaissances antérieures dans la langue nouvelle. Il n'est pas surprenant alors de voir que les élèves sont en mesure de faire la distinction entre plusieurs langues et de reconnaître les spécificités de chacune, et par là d'instaurer cette distinction à des fins pratiques. Dans ce sens, l'alternance codique devient une ressource langagière qui sert parfois à « dépanner » l'apprenant dans sa vie scolaire. Nous constatons ensuite des variations en ce qui concerne les représentations que les élèves ont sur les frontières des langues. Pour apprendre l'anglais, les deux élèves ont recours à des langues différentes, qui ne sont pas forcément leurs langues premières ou secondes (*cf.* extraits A-C). Par ailleurs, nous pouvons dire que des liens effectifs s'établissent entre le choix de langue, à laquelle les élèves ont recours, et la personne avec laquelle ils parlent cette langue. Nous considérons, à ce titre, que, dans la construction de la compétence plurilingue chez les élèves, la représentation d'une proximité affective entre les langues est plutôt privilégiée au détriment d'une représentation objective de la distance typologique entre les langues. Il est difficile de déterminer, à ce stade, si les représentations influencent le processus d'acquisition des langues chez l'apprenant. En revanche, nous pouvons affirmer que les représentations contribuent à la construction progressive d'une conscience métalinguistique sur les liens entre les langues, comme nous le constatons par le biais des extraits analysés au long de cet article.

Enfin, ces quelques considérations nous amènent à dire que l'emploi de plusieurs langues dans l'interaction n'est pas un phénomène aléatoire, mais qu'il assure au contraire des fonctions pragmatiques dans l'élaboration du discours. Dans cette perspective, la manifestation des alternances codiques dans les pratiques langagières des apprenants relève plutôt des représentations affectives et subjectives qu'ont les apprenants sur leur propre compétence plurilingue (nous pouvons ici citer, à titre d'exemple, la sécurité ou l'insécurité quant à leurs compétences), de la compétence plurilingue de leur interlocuteur et des facteurs situationnels et interactionnels, ou plus largement sociolinguistiques. Les témoignages montrent que les locuteurs choisissent les langues lors de l'interaction, en procédant par des analyses logiques, mais aussi en prenant appui sur des liens affectifs. Ainsi, en analysant les points communs et les différences entre les langues, les élèves utilisent leurs connaissances comme des normes de référence pour l'acquisition de nouvelles compétences.

Un dernier point mérite d'être souligné à propos de la variété des fonctionnements linguistiques mis en évidence à travers les trois extraits. En nous penchant brièvement sur le rôle de la L1 dans l'apprentissage des langues, nous pouvons dire que l'emploi de la L1 (ou des L1) devient une étape de développement de l'interlangue chez l'apprenant et un pas vers l'acquisition de la compétence plurilingue. En d'autres termes, ces extraits nous montrent le rôle que peut jouer la confrontation avec divers systèmes linguistiques dans le développement d'une conscience métalinguistique chez l'apprenant. C'est dans ce sens que nous postulons qu'une réflexion systématique sur le bi-/plurilinguisme permettrait, d'une part, de déterminer avec précision le pouvoir de l'articulation des langues dans la régulation communicationnelle des interactions individuelles et, d'autre part, de mieux comprendre le rôle des représentations dans la construction de la compétence plurilingue chez l'apprenant.

Références bibliographiques

- AYMONOD, Pierangelo, CAVALLI, Marisa, COSTE, Daniel, DEMATTEIS, Fulvia, PORTE, Graziella, ROSINA, Maurizio, SCIACQUA, Cinzia (2006). *Langues, apprentissages, identité*. Val d'Aoste : IRRE-VDA.
- CONSEIL DE L'EUROPE (2001). *Cadre européen commun de référence pour les langues*, Paris : Didier.
- COSTE, Daniel (2004). De quelques déplacements opérés en didactique des langues par la notion de compétence plurilingue, in *Structures et discours. Mélanges offerts à Eddy Roulet*, Auchlin, Antoine *et al.* (Eds). Québec : Nota bene, 67-85.
- COSTE, Daniel, MOORE, Danièle, ZARATE Geneviève (1997). *Compétence plurilingue et pluriculturelle*, Strasbourg : Conseil de l'Europe.
- MOORE, Danièle (2006). *Plurilinguismes et école*, Paris : Didier.
- STRATILAKI, Sofia (2004a). Pluralité des représentations sociales et des pratiques bi-/plurilingues dans les lycées franco-allemands de Buc et de Sarrebruck, *Cahiers du français contemporain*, 9 : 173-190.
- STRATILAKI, Sofia (2004b). Mon chat, je lui parle en allemand même s'il vient de France, Contacts de langues et représentations du plurilinguisme chez les jeunes apprenants frontaliers, *REPÈRES*, 29 : 233-250.
- STRATILAKI, Sofia (2005). Vers une conception dynamique de la compétence plurilingue : quelques réflexions six ans après..., in *Plurilinguisme et apprentissages. Mélanges Daniel Coste, Mochet, Marianne. et al.* (Eds), Lyon : ENS Éditions, 155-168.
- STRATILAKI, Sofia (2006). Et si on jonglait un peu avec les langues ? Construction et gestion des répertoires plurilingues chez les apprenants franco-allemands, *Bulletin suisse de linguistique appliquée (VALS-ASLA)*, 83/1 : 121-132.
- STRATILAKI, Sofia (2007, à paraître). Composantes, structure opératoire et dynamique de la compétence plurilingue : modes d'articulation et formes de construction, in *Perspectives de la recherche francophone autour de la notion de compétence plurilingue*, Moore, Danièle & Castellotti, Véronique (Eds). Collection Transversales, Berne : Peter Lang.

LA PLACE ET LES FONCTIONS RÉSERVÉES À LA LANGUE MINORÉE : LE CAS DE BAFRA (GRECE)

Maria ZERVA

mzerva@umb.u-strasbg.fr

Maison inter-universitaire des Sciences de l'Homme-Alsace

Groupe d'Etude sur le Plurilinguisme Européen (GEPE), composante de l' EA 1339-Linguistique, Langues et Parole (LILPA)

Département d'Etudes Néo-helléniques

22 rue René Descartes

67084 Strasbourg Cedex

Abstract : A case of diglossia is marked by an unequal division of the functions and the uses of the present languages. This paper examines the situational context of the subordinate language use in a particular bilingual community, adding a discussion of the 'folklorisation' of the secondary language and the possibility of using this term in sociolinguistic research.

1. Objet de recherche et mise en contexte

Dans cet article, nous essaierons d'aborder la question des fonctions qui sont réservées à la langue minorée au sein d'une communauté bilingue. Il s'agit de Bafra, situé à 8 km de Ioannina, au nord-ouest de la Grèce. Cette communauté compte 670 habitants. Ils sont tous orthodoxes et bilingues (turc/ grec), d'où la particularité de cette communauté, à savoir l'association du turc et de l'orthodoxie chez les mêmes individus. Les Bafriotes sont issus d'Asie Mineure et se sont installés en Epire après l'échange mutuel des populations entre la Grèce et la Turquie, un échange convenu par le traité de Lausanne en 1923. Les premiers réfugiés étaient entièrement monolingues : ils ne parlaient que turc. Il est toutefois important de préciser ce que signifie « parler turc » dans le cas des Bafriotes. Ces derniers ne parlent pas le turc standard (du moins, ceux qui le parlent encore), mais une forme dialectale de turc, parlée à l'intérieur de l'Asie Mineure jusqu'aux années vingt, c'est à dire avant les réformes de Kemal Atatürk concernant la langue turque. De ce fait, le turc standard leur est aujourd'hui presque incompréhensible.

Le turc de Bafra peut être désigné aujourd'hui comme une langue fortement dominée par le grec, qui est la langue de prestige institutionnalisée et qui constitue un des éléments fondamentaux de l'identité grecque, d'après la façon dont elle est conçue par la politique. L'abandon du turc au profit du grec chez les locuteurs en question est évident et il en va de même pour la réduction de ses fonctions, comme nous le montrons dans notre mémoire de DEA (Zerva, 2005 : 96-107). Néanmoins, malgré le fait que le turc a perdu une grande partie de sa fonctionnalité au fil des générations (4 générations depuis l'installation des premiers réfugiés dans les années 20), il paraît être une langue toujours fortement présente dans la vie de la communauté, et cela à partir de quelques fonctions particulières qui lui sont confiées. La présence de cette langue est même attestée chez les plus jeunes Bafriotes ainsi que chez les gens qui ne le maîtrisent pas.

Dans cet article, nous tenterons de préciser quelles sont les fonctions qui résistent encore à la force assimilatrice du grec et nous nous efforcerons de les confronter, afin de pouvoir expliquer le maintien du turc. Pour ce faire, nous nous appuierons sur une enquête menée sur le terrain, réalisée au moyen d'entretiens semi-directifs sur la base d'un questionnaire spécialement conçu pour cette communauté et qui a fait l'objet d'un enregistrement audio de

neuf heures. Nous examinerons par la suite les discours recueillis en nous fondant sur l'analyse minutieuse de Zerva (2005 : 38-65). Notons que nous procéderons à l'analyse de nos données en opérant une répartition des interviewés en groupes selon leur sexe et leur âge, ce qui fait six sous-groupes (femmes et hommes de plus de 65 ans, de 41 à 64 ans (âge moyen) et de 20 à 40 ans).

2. Encadrement théorique

Nous n'allons pas insister sur l'aspect théorique. Nous travaillerons sur un exemple de diglossie, pris dans son acception dynamique (une acception éloignée des définitions de Ferguson et Fishman et inspirée plutôt de la reprise de la notion par les linguistes « natifs » qui l'ont considérée dans son aspect conflictuel), en nous appuyant notamment sur la définition de Blanchet (2000). Ainsi on pourrait postuler que dans toute situation diglossique les deux langues (ou variétés) qui coexistent se trouvent dans un rapport de force déséquilibré. Selon Blanchet (2000 : 130), *« on parle de diglossie quand deux langues (ou lectes), ou plus, sont en situation hiérarchique (et donc forcément conflictuelle à un certain degré), l'une étant socialement positionnée comme dominant l'autre, sur un plan ou sur un autre (principalement le prestige social), ce qui provoque une répartition des champs d'usages¹. Cette répartition fonctionne selon une contrainte sociale générale »*.

Ainsi, dans de telles situations, nous sommes témoins d'un processus de minoration. Toujours selon Blanchet (2000 : 131) *« la diglossie est un processus de minoration (péjoration ethno-socioculturelle) d'un groupe humain, notamment à travers sa langue. Cette dévalorisation s'accompagne souvent d'une minorisation : réduction numérique proportionnelle ou brute des locuteurs par annexion dans un groupe alloglotte plus nombreux et/ou diminution du nombre d'occurrences et de locuteurs de la langue en question à l'intérieur de son groupe originel, ce qui signifie qu'une partie des membres du groupe est amenée à adopter la langue dominante majoritaire, ponctuellement ou largement. La minoration est qualitative, elle joue sur le statut. La minorisation est quantitative, elle joue sur les pratiques¹. L'addition de ces deux processus liés conduit le groupe ethno-socioculturel minoré et minorisé à la situation de groupe (de langue) minoritaire »*.

Il n'est pas moins vrai que tout processus de minoration engendre un processus de majoration, ce qui fait qu'une langue (une variété ou un dialecte) n'est jamais évaluée entièrement de façon négative ou inversement. Ainsi, *« il n'y a pas de minoration/minorisation (= minoritarisation) sans majoration/majorisation (= majoritarisation) et réciproquement, à la fois dans l'ensemble d'un processus social¹ (la minoritarisation d'un groupe ou d'une pratique signifie en miroir la majoritarisation d'autres groupes ou d'autres pratiques) et à l'intérieur même du groupe ou de la pratique minoritarisée¹ (des aspects positifs sont associés aux aspects négatifs : un « patois » évalué négativement par rapport au français reçoit par exemple des attributs sociaux positifs de convivialité ou de nostalgie). Inversement, une langue majoritarisée fait aussi l'objet d'attributs négatifs [...] »* (Blanchet, 2005 : 34).

C'est dans ce sens que Gardy et Lafont (1981 : 77) parlent des *« représentations de la diglossie »*. Ces représentations suscitent selon eux une dévalorisation de la langue dominée, qui est d'ailleurs confirmée dans la pratique. Néanmoins, il est intéressant de constater que, parallèlement à cette dévalorisation, un autre processus s'opère *« dans un mouvement de compensation »* (*ibid*) : plus les locuteurs de la langue dominée se sentent menacés par la force hégémonique de la langue dominante, plus ils créent autour de la première une mythologie flatteuse qui lui accorde des qualités refusées à la langue dominante (beauté, harmonie, intimité, chaleur, proximité, solidarité...). Pourtant, cette *idéalisation* et

¹ Nous soulignons.

folklorisation de la langue dominée ne fait que conforter la position de la langue dominante, dont l'usage n'est pas une valeur, mais un fait (Gardy & Lafont, 1981 : 77 ; Boyer, 1996 : 20).

Dans le présent travail, nous nous proposons d'étudier les fonctions qui sont assignées à la langue minorée dans le cas précis de la communauté de Bafra, le turc. La réflexion sur les usages qui lui sont pratiquement confiés, nous amènera sans doute à revoir la notion de folklorisation de la langue dominée. Cette notion, peu utilisée dans la littérature sociolinguistique, nous a paru intéressante surtout par rapport à notre terrain d'étude et aux résultats de notre enquête. Il serait en effet légitime de s'interroger sur son sens, sur son éventuel apport au sein de la sociolinguistique et, enfin, sur la pertinence d'une telle notion.

3. La fonctionnalité des langues

Dans cette partie de notre travail, nous nous pencherons sur les discours de nos interlocuteurs concernant la maîtrise du turc, mais surtout son utilisation, à savoir la place qu'il occupe dans leur répertoire personnel. Pour ce faire, nous répartissons, comme déjà dit, nos interlocuteurs en sous-groupes selon leur âge et leur sexe.

3.1. La maîtrise de la langue minorée (turc)

Tout d'abord, il faut souligner que tous les interviewés, même les plus âgés², maîtrisent d'une façon satisfaisante le grec, ce qui nous a permis de réaliser la totalité des entretiens en grec ; et il est intéressant de prendre en compte cette donnée pour mesurer l'ampleur de la substitution linguistique qui a eu lieu durant les dernières décennies à Bafra, où la totalité des premiers réfugiés était initialement monolingue (turcophone), à de rares exceptions près.

Dans le cas des femmes, nous constatons un recul considérable de la connaissance du turc au fil des générations. Les plus âgées ont appris le turc en premier, mais ce n'est plus le cas dans les générations suivantes ; les femmes des deux dernières générations ont appris le turc essentiellement pour communiquer avec les grands-parents ou les beaux-parents. On voit donc le rôle des générations aînées dans le maintien du turc. Pour les femmes d'âge moyen, le turc est encore présent dans leur vie, même si elles maîtrisent mieux le grec. Néanmoins, la connaissance du turc s'affaiblit fortement chez les plus jeunes, qui en ont une connaissance plutôt passive.

En ce qui concerne les hommes, on constate également le recul du turc au fil des générations, mais il est moins prononcé que celui constaté chez les femmes. Pour les aînés, le turc est leur première langue ; il occupe également une place importante dans la vie des plus jeunes. En effet, tous les hommes interviewés ont une bonne connaissance du turc (à l'exception d'un seul) et l'ont appris dès leur enfance. Ils maîtrisent aussi le grec (fait constaté lors de l'enquête) et, si l'on prend en compte les résultats scolaires, on peut conclure que les deux derniers sous-groupes ont appris le grec très vite. Notons aussi que le rôle des beaux-parents dans l'apprentissage du turc par les femmes, surtout d'âge moyen, n'est pas valable pour les hommes, puisqu'ils apprennent le turc dans le foyer paternel (Zerva, 2005 : 99).

3.2. Les fonctions de la langue minorée

Pour ce qui est des fonctions des deux langues en présence, il est clair que pour les femmes, le turc perd sa fonctionnalité au fil des générations ; les plus âgées utilisaient le turc dans le foyer parental et l'utilisent même aujourd'hui pour communiquer avec leurs proches, que ce soit avec leur conjoint ou leurs amies. Il est pourtant remarquable que le turc n'est en aucun cas délibérément transmis à leur progéniture ; les femmes parlent toujours en grec à leurs

² Notons que les plus âgés de nos interlocuteurs appartiennent déjà à la deuxième génération des réfugiés.

enfants. Le turc perd en grande partie sa force communicationnelle à partir du sous-groupe d'âge moyen (femmes de 41 à 64 ans) et en est complètement dépourvu avec le dernier sous-groupe (femmes de 20 à 40 ans). Notons cependant que le turc est la langue des chansons préférées de toutes les interviewées, et que les usages principaux du turc auxquels elles font référence sont les danses, les chants, les lamentations, à l'occasion des mariages, des fêtes, des funérailles. Toutes les interviewées ont déclaré aimer la musique turque. Toutes les femmes âgées l'ont choisie comme leur musique préférée et il en va de même pour trois des quatre femmes d'âge moyen (une sur quatre a déclaré l'aimer mais préférer la musique et la danse pontiques). Chez les plus jeunes, on mentionne les deux types de musique ; une sur quatre préfère la musique avec des paroles turques, une choisit la musique grecque tandis que les deux autres aiment autant l'une que l'autre. C'est ce côté de l'utilisation du turc que nous appelons la « dimension folklorique » de cette langue, qui, à notre avis, est spécifique des situations de domination linguistique, où les locuteurs de la langue minorée, qui est dénuée des fonctions réelles, cherchent à lui accorder des qualités et des fonctions qui relèvent plutôt du domaine affectif. Ajoutons à cela qu'une partie de nos interviewées aime écouter le turc à la télévision, surtout la musique. Ce point va être aussi mis en relief lors de l'étude des discours des hommes. En outre, la fonction la plus couramment réservée au turc ces derniers temps est la conversation secrète et, moins mentionnés, la narration de plaisanteries et l'usage d'expressions figées. Ainsi, une de nos interlocutrices, qui a 76 ans, nous dit à propos de l'alternance codique :

I 76 : [On alterne les langues] surtout quand il y a un groupe d'amis et qu'on veut discuter, dire des plaisanteries ou des blagues, on parle en grec et aussi en turc.

E : C'est-à-dire que les blagues [sont] surtout en turc ?

I 76 : Surtout en turc. Surtout. Euh, c'est quand même notre langue maternelle...

La nature des expressions figées n'est pas facile à préciser. A ce propos, la plus jeune interviewée (25 ans) nous dit :

I f 25 : On peut le parler [le turc]. Mais, en général, même si je suis avec d'autres, j'ai quelques phrases standard que j'utilise quand quelque chose de précis m'arrive.

E : C'est quoi à peu près, c'est, bon, des injures, je ne sais pas, moi...

I f 25 : Non, pas d'injures. Je ne sais pas, moi...

Tierce personne (30 ans) : Si on veut dire quelque chose en cachette des enfants, se comprendre entre nous, on l'utilise.

I f 25 : Ou, disons, quand, comme ça, on en a assez (petit rire), on l'utilise, comme ça, quelques, euh, phrases standard. Voilà.

Il est intéressant de constater que 64% de nos interlocutrices se réfèrent au choix du turc quand elles ne veulent pas être comprises par d'autres personnes présentes (une sur trois des femmes âgées, quatre sur quatre de celles d'âge moyen et deux sur quatre des plus jeunes). Le taux bas du premier sous-groupe pourrait s'expliquer, à notre avis, par l'usage généralisé du turc par cette génération.

Pour les hommes aussi, le turc perd sa fonctionnalité au fil des générations et devient une langue « accessoire » pour les nouvelles générations ; le refus de transmettre le turc à la progéniture en est une preuve³. Pour les plus âgés, le turc a servi de langue de

³ Ce refus paraît contradictoire avec la maîtrise du turc par la quasi-totalité des hommes ; comme les hommes l'utilisent en général plus entre eux, il nous semble que c'est plutôt par la socialisation

communication dans le foyer parental et ce n'est que la scolarisation qui a introduit le grec dans leur répertoire linguistique. Le grec est pour eux la langue utilisée avec le monde extérieur et les nouvelles générations ; avec les proches, les personnes du même âge et du même environnement spatio-culturel (à savoir des réfugiés turcophones), cette génération utilise le turc encore aujourd'hui. C'est à partir de la génération d'âge moyen que le grec devient la langue principale, qu'on utilise au quotidien et le turc sert surtout à communiquer avec les plus âgés, à parler en secret, à dire des injures, à raconter des histoires ou des proverbes transmis par les anciens, ou pour de petits morceaux dialogaux (du type « bonjour », « ça va »), pour que le locuteur affirme son identité distincte et souligne les liens de proximité avec son interlocuteur. Les fonctions le plus souvent mentionnées sont la communication en secret (signalée par 35,7% des interviewés), les injures (35,7%) et les plaisanteries (28,5%). Pour donner quelques exemples :

- *I h 56 : [On parle turc] euh, euh, de temps en temps, comme ça, pour rigoler, si on veut dire quelque chose, comme ça, un secret, voilà.*
- *I h 38 : Si je me trouve avec des gens de ma bande, de mon âge et tout ça, il se peut que je me mette à parler turc. Pour rigoler, pour des femmes, pour je ne sais pas quoi, ça peut arriver, bien sûr.*

E : Si des étrangers, des natifs sont présents ?

I h 38 : Non, là je ne parle pas, je ne parle pas. Seulement s'il faut dire des gros mots ! (rires) Euh, XXX disons les choses comme elles sont !

Le turc a encore une place importante dans le répertoire des hommes de plus de 60 ans, mais reste une langue proche, et même intime à tous ; ainsi, quelques-uns d'entre eux, regardent la télévision turque ou ont l'intention de le faire dans un futur plus ou moins proche, et entretiennent une relation particulière avec la musique « aux paroles turques ». Il nous semble plus adéquate de la nommer ainsi puisque, même dans la conscience des Bafriotes, ce ne sont que les paroles qui sont en turc ; la musique – la mélodie – est orientale et pourrait être considérée aussi bien comme turque que comme grecque, puisque la musique des Grecs d'Asie Mineure avait bien sûr des points communs avec la musique de leurs voisins Turcs⁴. En effet, un peu plus de la moitié (57,1%) des interviewés préfèrent la musique aux paroles turques (8 sur 14, dont la moitié n'a mentionné que la musique turque et le reste l'a mise en parallèle avec la musique grecque) ; trois sur quatorze (21,4%) aiment la musique turque autant que la musique grecque, tandis que deux sur quatorze (14,2%) déclarent préférer la musique grecque et un seul interviewé (7,1%) prétend aimer plus la musique grecque, mais nous nous permettons de contester son affirmation, vu l'ensemble de son discours particulièrement surveillé et parfois contradictoire. Il faut souligner aussi que quatre hommes d'âge moyen sur cinq (28,5% dans l'ensemble et 80% dans ce sous-groupe) se réfèrent au côté affectif de la musique aux paroles turques, et cela indépendamment de leur préférence (grecque ou turque) ; de la sorte, la musique turque est celle qui « les touche », qui les émeut (surtout parce qu'elle fait appel aux origines lointaines et aux parents).

I h 63 : Euh, je vous dirais... autant l'un que l'autre. Les [chansons] grecques et aussi les... ça me touche beaucoup les... karsilamas⁵. En général, moi, je peux même paraître ridicule quand j'entends les... les nôtres. [...] Elles ont un ton plaintif... On les tient de nos parents, ça me fait... Euh... Parce que je sais qu'ils ont beaucoup souffert. Toutes les chansons ont un sens, ces chansons-là.

« masculine » (par exemple dans des lieux de haute fréquentation masculine, comme le *kafeneio* – la cafétéria – du village) que par la transmission délibérée que les jeunes apprennent le turc.

⁴ Notons que deux de nos interlocuteurs l'ont appelée « musique orientale de provenance byzantine ».

⁵ Mélodie orientale.

*Elles étaient très douloureuses, toujours avec des plaintes, avec détresse...
Nos parents ont beaucoup souffert...*

Globalement, le turc est plus utilisé par les hommes, même les plus jeunes, que par les femmes (cf. 3.1) ; l'alternance codique, pratiquée même par les plus jeunes hommes, contrairement aux jeunes femmes, en est une preuve (Zerva, 2005 : 102-106).

Résumons en quelques mots nos observations. A Bafra, le processus d'assimilation linguistique est évident ; le turc parlé par ses premiers habitants dans les années vingt a perdu une grande partie de sa force communicationnelle, dès lors qu'il a été remplacé par le grec. Il est intéressant que cela soit arrivé sans résistance, presque comme une nécessité, pour que les Bafriotes puissent s'intégrer dans la société grecque en tant que membres égaux. En abandonnant le turc, les réfugiés turcophones de la Grèce, et non pas seulement ceux de Bafra, ont montré leur détermination à faire partie de la large communauté hellénique et à s'approprier entièrement l'identité grecque (qui ne leur était d'ailleurs pas accordée volontiers par les populations natives) dans tous ses aspects, et notamment en ce qui concerne la langue, élément identitaire prépondérant. Ainsi, le grec, qui par le passé a été introduit dans la communauté par le seul moyen de la scolarisation, est aujourd'hui la langue quotidienne de la plupart des Bafriotes, surtout ceux âgés de moins de 60 ans ; c'est la langue de la scolarisation, du travail, de l'église, des relations interpersonnelles. Le turc n'a aucune existence officielle ou publique et reste une langue strictement réservée à la communauté. Néanmoins, le turc, bien qu'il ait perdu son rôle prééminent dans la société bafriote, y est toujours présent, ce qui est surtout manifeste dans le cas des hommes, comme on l'a vu en détail dans ce qui a précédé.

3.3. La folklorisation : essai de modélisation

Un aspect particulièrement intéressant du maintien du turc chez les Bafriotes est le rôle de la musique qui constitue, à notre avis, un phénomène plus large et une question à approfondir. Ce sujet de la musique préférée fait presque l'unanimité : la musique orientale avec ses paroles en turc est de loin celle qui touche le plus les Bafriotes, celle qui est considérée comme la plus intime et, en un mot, comme « leur musique à eux ». Il est intéressant de mentionner aussi qu'en 2001, l'association culturelle de Bafra a sorti un CD intitulé « Danses et chansons traditionnelles de Bafra Ioannina » qui ne comprenait que des chansons aux paroles turques⁶. Ce rapport à la musique qui constitue un trait saillant du rapport que la communauté entretient avec le turc, nous a amenée à réfléchir sur l'éventuelle utilisation d'une notion que l'on pourrait désigner par le nom de « folklorisation » de la langue minorée.

Comme nous l'avons mentionné plus haut, Gardy et Lafont (1981 : 77) parlent d'une *idéalisée* de la langue en position faible et Boyer (1996 : 16 ; 2001 : 53) y ajoute une *folklorisation* de la langue minorée, sans toutefois préciser le contenu exact que recouvre ce terme. Blanchet (2005) suggère d'ailleurs que toute minoration va de pair avec une majoration. Selon les données résultant de notre propre enquête, le turc en tant que langue en situation faible, se trouve donc en grande partie dépourvu de fonctions réelles (donc minoré), mais est toutefois pratiqué en maintes occasions bien précises, parmi lesquelles la musique est la plus saillante. Ainsi, il nous paraît important de discuter l'usage de la notion de folklorisation. Or, des questions qui demandent des réponses convaincantes surgissent ; une telle notion serait-elle pertinente ? Est-ce une notion que l'on pourrait retenir au sein de la sociolinguistique ? S'agit-il d'une notion qui décrirait une fonction parmi d'autres de la

⁶ Danses et chansons traditionnelles de Bafra Ioannina (en grec) (2001). Association culturelle de Bafra Ioannina « Sainte Paraskevi » et Association Internationale d'Art Populaire.

langue minorée ou bien l'utiliserait-on pour des situations où cette langue ne survit que par la folklorisation?

Rien de simple dans cette tentative, néanmoins nous relèverons le défi et nous continuerons dans le but de trouver un cadre théorique convenable pour notre étude de cas. Ainsi, selon des dictionnaires d'ethnologie et d'anthropologie (Bonte et Izard, 2002 : 283-4 ; Pluriel recherches, 2001 : 23-25), le *folklore* renvoie étymologiquement au « savoir traditionnel pratique » du peuple (Pluriel recherches, 2001 : 23). Il désigne, en effet, « les contes et légendes, les croyances et pratiques religieuses, les coutumes, les savoirs empiriques, etc., de tradition en milieu rural [...] et en même temps, l'étude de ces phénomènes » (Pluriel recherches, 2001 : 23). Il est vrai qu'à un moment le mot a été doté d'une connotation péjorative (Pluriel recherches, 2001 : 25) et c'est pour cette raison qu'il a été en grande partie abandonné en sciences humaines. En même temps, il s'est étendu aux savoirs empiriques et aux pratiques d'autres milieux que les paysans, à savoir aux « milieux populaires citadins, immigrés, toutes les cultures ou sous-cultures minoritaires en général » (Pluriel recherches, 2001 : 25). Folkloriser serait *traiter comme un folklore*, d'où le nom folklorisation.

A notre connaissance, la *folklorisation* ne constitue pas une notion largement utilisée en sciences humaines et, plus particulièrement en sociolinguistique, d'autant plus que la signification de cette notion n'est pas claire. Fishman (1989 : 394-6, 400) s'en sert pour désigner la restriction des fonctions attribuées à une « langue en danger » aux univers les plus intimes de l'individu (famille, voisins, échanges de la vie quotidienne), tout en faisant allusion à l'aspect péjoratif du mot ; ainsi, ce rétrécissement de sens peut être perçu comme une banalisation de la langue en question au moyen d'une focalisation sur le « passé rural (contes des vieilles femmes, chansons folkloriques et danses d'un temps passé) » (*ibid.* : 394, notre traduction). Cet usage fait par *Fishman*, loin d'être explicite, nous fournit tout de même une matière première pour travailler sur notre problématique.

Pour répondre alors à notre questionnement, nous pensons que nous pourrions tirer profit de l'usage de cette notion à condition de le faire avec prudence, étant donné la connotation négative du mot *folklore*. Ainsi, notons premièrement que nous appréhendons la « dimension folklorique » dans son sens neutre, sans aucune connotation péjorative. La *folklorisation* serait pour nous une manifestation qui relève de la *fonction affective* de la langue, fonction familière aux sociolinguistes qui comprend tous les usages de la langue minorée ayant trait au monde intime de l'individu (usages au sein de la famille, des cercles amicaux, et en général tout échange non officiel et relatif aux domaines privés). Ce sont ces usages qui persistent le plus dans une situation de diglossie et de substitution linguistique avancée, puisque la langue en situation faible perd du terrain d'abord au niveau des fonctions publiques et officielles⁷.

La *folklorisation* a d'ailleurs un caractère plutôt *symbolique*, et désigne plus particulièrement la relation qu'une communauté entretient avec la musique et la danse à travers la langue minorée. Nous la considérons comme une *manifestation de la fonction affective parmi d'autres*, sans négliger le fait qu'elle est d'autant plus manifeste que le rétrécissement des fonctions de la langue minorée est patent. Ainsi, la *folklorisation* équivaut à une restriction des fonctions de la langue minorée, elle constitue donc un indice de minoration. Elle est pourtant accompagnée par une *idéalisierung* de cette même langue, qui constitue justement la majoration de celle-ci. Dans le cas des Bafriotes, le turc est donc en quelque sorte idéalisé à travers la musique, parce que cet usage fait appel aux origines lointaines de l'Asie Mineure et apporte ainsi une richesse identitaire aux Bafriotes, d'autant plus qu'ils considèrent cette culture comme supérieure à celle de la région d'accueil (Zerva, 2005 : 77-79, 116) ; le turc

⁷ Même si on ne peut pas être sûr du statut du turc de Bafra sur le territoire ottoman (avant 1923), on pourrait supposer qu'il était muni de ces fonctions, étant une variante de la langue de l'administration.

devient un moyen pour manifester leur positionnement identitaire particulier et en même temps un moyen qui les rapproche de leurs ancêtres et de leurs parents. Ce rapport *folklorisation* - *idéalisation* renvoie au couple *minoration* - *majoration* dans le sens de Philippe Blanchet, selon lequel la minoration va de pair avec un processus de majoration. Ainsi, les locuteurs de la langue minorée, bilingues, étant donné l'avancement du processus d'assimilation linguistique, lui attribuent les qualités positives d'intimité, convivialité, etc. pour compenser sa dévalorisation au niveau des usages et atténuer ainsi la tension qui en résulte au niveau des représentations des langues et ainsi de leurs positionnements identitaires.

4. Conclusion

Dans cet article, nous avons tenté de circonscrire les fonctions qui sont réservées à la langue minorée dans le cas de la diglossie (grec-turc) de Bafra. Des résultats de l'analyse, il ressort que la langue en situation faible, à savoir le turc, survit aujourd'hui chez les nouvelles générations surtout en tant que code communicationnel secret et langue de la musique préférée. Nous nous sommes intéressée plus particulièrement à ce dernier aspect et nous avons discuté la possibilité d'utiliser le terme de *folklorisation* pour désigner ce rapport à la musique en tant que manifestation de la fonction affective de la langue minorée. Ainsi, nous appréhendons la folklorisation comme restriction, et partant, comme minoration ; il n'empêche qu'elle entraîne une idéalisation, à savoir une majoration de la langue en question. Ainsi, il nous semble que la notion de folklorisation ainsi retravaillée pourrait être fort utile pour les études de terrain, à vocation sociolinguistique où l'on observe des attitudes semblables au sujet de la musique, surtout quand cette dernière joue un rôle déterminant dans la préservation de la langue en position faible.

Références bibliographiques

- BLANCHET, Philippe (2000). *La linguistique de terrain. Méthode et théorie. Une approche ethno-sociolinguistique*. Rennes : Presses Universitaires de Rennes.
- BLANCHET, Philippe (2005). Minorations, minorisations, minorités : Essai de théorisation d'un processus complexe, in *Cahiers de Sociolinguistique* n° 10, *Minorations, minorisations, minorités. Etudes exploratoires*, Huck, Dominique & Blanchet, Philippe (dirs). Rennes : Presses Universitaires de Rennes, 17-47.
- BONTE, Pierre & IZARD, Michel (2000). *Dictionnaire de l'ethnologie et de l'anthropologie*. Paris : Quadrige/ PUF.
- BOYER, Henri (dir) (1996). *Sociolinguistique : Territoire et objets*. Lausanne : Delachaux et Niestlé.
- BOYER, Henri (2001). *Introduction à la sociolinguistique*. Paris : Dunod.
- GARDY, Philippe & LAFONT, Robert (1981). La diglossie comme conflit : l'exemple occitan, *Langages*, 61: 75-91.
- FISHMAN, Joshua A. (1989). *Language and ethnicity in minority sociolinguistic perspective*. Clevedon- Philadelphia : Multilingual Matters.
- Pluriel recherches n° 8 (2001). Vocabulaire historique et critique des relations inter-ethniques. Paris : L'Harmattan (entrée folklore).
- ZERVA, Maria (2005). *Langues en contact et hégémonie linguistique : Le cas des Grecs orthodoxes turcophones de Bafra (Ioannina, Grèce)*. Mémoire de D.E.A. Département d'Etudes Néo-helléniques, Université Marc Bloch, Strasbourg (inédit).

DIDACTIQUE DES LANGUES

CARACTÉRISATION DE LA NOTION D'ERREUR LEXICALE ET DESCRIPTION DES PROBLÈMES LEXICAUX D'ÉTUDIANTS UNIVERSITAIRES

Dominic ANCTIL

dominic.anctil@umontreal.ca

Département de didactique, Université de Montréal

Observatoire de linguistique Sens-Texte (OLST)

Canada

Abstract : This paper aims to characterize the notion of lexical error, which has not been defined in either language teaching or in linguistics. Within the framework of Meaning-Text Theory, the different properties of the lexical units are looked at, as well as the ways in which they could be violated to occasion lexical errors. This provides all the necessary elements to describe lexical problems found in a text corpus.

Introduction

Les quelques recherches s'étant intéressées à la maîtrise du français écrit chez les étudiants aux ordres supérieurs d'enseignement (par exemple Lefrançois, 2004; Roy, Lafontaine et Legros, 1995) n'ont révélé que peu d'informations sur leur maîtrise lexicale. Ces travaux portant principalement sur l'orthographe grammaticale, l'aspect vocabulaire y est traité de façon plutôt superficielle. Malgré ce portrait peu détaillé de la dimension lexicale de l'écrit des étudiants, ces recherches ont tout de même révélé des informations intéressantes : 1) les problèmes lexicaux sont présents dans les rédactions d'étudiants, 2) le lexique est la dimension la moins prise en compte par les étudiants en révision/rédaction (Roy *et al.*, 1995) et 3) les erreurs lexicales sont les moins repérées et les moins souvent résolues correctement (Lefrançois, 2004). Par contre, d'un point de vue didactique, une description plus fine des problèmes lexicaux des étudiants s'avère nécessaire, puisqu'elle constitue une première étape vers des interventions visant à les corriger.

Le manque de descriptions détaillées des erreurs lexicales des étudiants découle peut-être d'un problème plus profond : l'absence de consensus autour de la notion même d'erreur lexicale. En effet, nombreux sont les auteurs à souligner la difficulté de cerner l'erreur lexicale (Lépine, 1995; Patris et Vansnick, 1992; Polguère, 2004); Denance (1999) pointe la difficulté de départager les erreurs de syntaxe des erreurs de vocabulaire, alors qu'Engber (1995) mentionne le caractère subjectif des choix lexicaux en rédaction. Il semble que ni la didactique ni la linguistique n'aient jusqu'à maintenant proposé une caractérisation claire et complète de ce qu'est une erreur lexicale. Pourtant, les avancées importantes de la linguistique en ce qui a trait à la description lexicale et sémantique nous ont fourni tous les éléments conceptuels nécessaires à une telle caractérisation.

Voici donc les deux objectifs que vise le présent article :

1. Proposer une caractérisation claire et exhaustive de l'erreur lexicale appuyée par des concepts linguistiques.
2. Décrire, à la lumière de la définition proposée dans le premier objectif, les problèmes lexicaux d'étudiants universitaires.

2. Le cadre théorique

Il est important que la caractérisation de l'erreur lexicale que nous proposons s'appuie sur des concepts linguistiques clairs. Le cadre théorique que nous avons choisi est la Théorie Sens-Texte (TST) (Mel'čuk, 1997), qui place le sens au cœur de sa modélisation de la langue. Le lexique en étant le principal élément porteur de sens, la description lexicale occupe une place de choix dans la TST. En conséquence, la Lexicologie Explicative et Combinatoire (LEC) (Mel'čuk, Clas et Polguère, 1995), branche lexicologique de la TST, s'est énormément développée et offre un cadre idéal pour l'élaboration d'une définition de l'erreur lexicale.

La TST considère les lexies comme des entités à trois dimensions. En plus de la classique association signifiant-signifié proposée par Saussure (1960), la TST prend en compte une troisième dimension : celle des propriétés de combinatoire, qui regroupent l'ensemble des informations que l'on doit connaître sur un signe linguistique, en plus de sa forme et de son sens, pour pouvoir l'utiliser correctement (voir 3.2.3.). Évidemment, une telle conception a une influence sur l'approche lexicologique de la LEC, qui doit aussi considérer cette troisième dimension des lexies dans leur description. La LEC vise donc, comme toutes les approches lexicologiques, à décrire le sens des lexies; c'est sa dimension explicative (voir 3.2.2.). Par contre, un deuxième axe de description est plus typique de la LEC : la description des différentes propriétés qui régissent le comportement syntaxique et lexical des lexies. C'est la dimension combinatoire de la LEC. En d'autres mots, la LEC vise à fournir un portrait exhaustif de tout ce que doit savoir un locuteur pour pouvoir utiliser correctement une lexie. Une connaissance aussi complète des propriétés en jeu dans la maîtrise d'une unité lexicale sera bien sûr très précieuse lorsque viendra le temps de caractériser la notion d'erreur lexicale.

3. La notion d'erreur lexicale

Cette section vise la caractérisation de la notion d'erreur lexicale. Par la présentation des diverses normes régissant l'emploi du lexique et l'exploration des différentes dimensions de l'unité lexicale, nous établirons une typologie des problèmes lexicaux.

3.1. L'erreur lexicale et la norme

Nous appelons communément « erreur » un usage qui est déviant par rapport à une norme; il est donc facile de pointer une erreur tant et aussi longtemps que nous disposons d'une norme bien définie. La grammaire scolaire, d'ailleurs souvent qualifiée de normative, constitue un bon exemple de norme bien définie. Par contre, il existe des dimensions du code linguistique pour lesquelles la norme ne semble pas aussi clairement définie. Le lexique en est un bon exemple; si certaines dimensions de l'emploi du vocabulaire semblent soumises à une norme rigide – par exemple l'orthographe lexicale –, il en existe d'autres pour lesquelles la norme est plus floue – pensons notamment à la collocation (voir 3.2.3.). Par conséquent, il nous faudra, pour caractériser la notion d'erreur lexicale, recourir à plusieurs types de normes.

3. **La norme prescriptive** : Souvent appelée le « bon usage », cette norme domine la grammaire scolaire. Il s'agit d'un « ensemble de recommandations déterminées par une partie de la société et précisant ce qui doit être retenu parmi les usages d'une langue afin d'obtenir un certain idéal esthétique ou socioculturel. » (Legendre, 2005: 938) Certains organismes, comme l'Office québécois de la langue française ou l'Académie française, ont d'ailleurs pour mission de fixer ce « bon usage ». Si cette norme vise surtout les aspects grammaticaux, elle exerce aussi son emprise sur certains aspects du lexique comme l'orthographe d'usage, le régime verbal (fait pour un verbe d'être transitif direct/indirect ou non), le genre nominal, l'invariabilité, etc.
4. **La norme descriptive** : Il n'existe pas beaucoup de « règles lexicales » aussi rigides que celles régissant les emplois grammaticaux; par conséquent, la norme prescriptive ne

saurait suffire à caractériser l'ensemble des erreurs lexicales. En effet, certains emplois lexicaux pourront être perçus comme douteux, non pas en vertu d'une règle spécifique, mais plutôt parce qu'ils ne correspondent pas à un usage « normal », c'est-à-dire qui n'est pas marqué par rapport à la façon dont les locuteurs s'expriment en général. Cette vision de la norme est plutôt descriptive et constitue le point de vue communément adopté par les linguistes : la langue est décrite telle qu'elle est réellement utilisée par les locuteurs et l'usage de la majorité devient la norme. Pour juger de l'acceptabilité d'un emploi lexical du point de vue descriptif, nous pourrions avoir recours au jugement de locuteurs natifs ou utiliser des corpus pour en vérifier la fréquence.

5. **La norme fonctionnelle** : L'importance de la prise en compte des dimensions pragmatiques dans la communication est aujourd'hui généralement admise. Il se peut donc que certains emplois lexicaux soient perçus comme problématiques parce qu'ils entrent en conflit avec le contexte de communication (ex. registre de langue inapproprié). Pour diagnostiquer de tels problèmes, nous aurons recours à une norme fonctionnelle, selon laquelle « le bon usage est celui qui est adapté aux diverses situations de communication, aux situations de la vie en société » (Dallaire, 1997: 30).
6. **La norme stylistique** : La variété et la richesse du vocabulaire constituent deux points d'évaluation très communs dans les grilles de correction de textes. Pourtant, rarement des critères précis servant à les opérationnaliser sont fournis. Par conséquent, l'évaluation en est la plupart du temps faite de façon plutôt subjective; ces aspects du vocabulaire sont soumis à une norme floue, que nous pourrions qualifier de « stylistique ».

3.2. *Élaboration d'une typologie des problèmes lexicaux*

Voici comment nous pourrions définir une erreur lexicale.

Erreur lexicale : Résultat de la transgression de la norme concernant une propriété d'une unité lexicale (sens, forme, propriété de combinatoire)

Pour proposer une typologie complète des différents types de problèmes lexicaux, il faudra donc observer quelles sont les transgressions possibles pour chacune des propriétés des unités lexicales; il faudra pour ce faire déterminer quelle norme s'applique à chacune. Les sections suivantes proposent donc un tour d'horizon des différentes dimensions des lexies et des erreurs lexicales qui y sont liées.

3.2.1. Le signifiant

Les erreurs lexicales relatives à la forme ayant déjà fait l'objet de plusieurs travaux (notamment Catach, Gruaz et Duprez, 1986), elles présentaient pour nous un intérêt moindre. La forme étant certainement la dimension des unités lexicales la plus fortement soumise à la norme prescriptive, il est facile d'imaginer de quelle façon cette dimension peut être transgressée : en ne respectant pas l'orthographe recommandée. Nous pouvons quand même différencier quelques types d'erreurs relatives à la forme.

Les erreurs d'orthographe lexicale sont celles qui découlent d'une mauvaise transposition phonème-graphème, mais sans être perceptibles à l'oral. Par exemple, **échalotte*, *aller de *paire (avec)* et **quoiqu'il en soit* présentent des erreurs d'orthographe lexicale. Les barbarismes, quant à eux, témoignent d'une méconnaissance plus profonde du signifiant d'une lexie et sont perceptibles même à l'oral, comme **assuidité* ou une **bride de conversation*. Évidemment, différents phénomènes sont en jeu (inversion de phonèmes, ajout de syllabe, etc.) et une caractérisation plus fine des barbarismes serait possible. Nous pouvons aussi étendre les barbarismes aux locutions déformées – par exemple *dire *l'inverse* (→ *une chose*) et *son contraire*, à la **lueur* (→ *lumière*) *de ces informations* –, puisqu'il s'agit encore une fois d'une méconnaissance manifeste de la forme d'une lexie. Finalement, le dernier cas de transgression de la forme est la production d'une forme inexistante, soit par

un emprunt à une autre langue (ex. **marchandiser* → *to merchandise*) ou par une mauvaise dérivation morphologique (ex. *un *courseur automobile*).

3.2.2. Le signifié

Pour bien comprendre les types d'erreurs lexicales liées au sens des unités lexicales, il s'impose d'établir clairement quelles sont les composantes d'une définition. Pour ce faire, observons la définition du verbe RAMPER [*Les soldats rampèrent sous la barrière.*] tirée du dictionnaire Antidote (2006), mais qui correspond aux méthodes de définition de la LEC.

rampier : Progresser lentement ventre contre le sol, en s'aidant de ses quatre membres, en parlant d'animaux, de l'homme.

Une telle définition consiste en une paraphrase, en termes sémantiquement plus simples, de la lexie décrite. Il s'agit d'une définition de type analytique, puisqu'elle comporte deux parties : le genre prochain et les différences spécifiques. Le genre prochain d'une définition en est l'élément central; il s'agit du noyau sémantique du sens décrit. La plupart du temps, le genre prochain est un hyperonyme de la lexie décrite. Par exemple, le genre prochain de 'rampier' est ici 'progresser'. Les différences spécifiques, quant à elles, sont des composantes sémantiques périphériques qui viennent restreindre, ou spécifier, le sens du genre prochain de façon à le faire correspondre au sens décrit et de différencier ce dernier d'autres sens qui partageraient le même genre prochain. Dans ce cas-ci, les différences spécifiques sont 'lentement', 'le ventre contre le sol' et 'en s'aidant de ses quatre membres'.

Pour certains types d'unités lexicales, dites prédicatives, il faudra aussi ajouter des informations d'une autre nature : le typage des participants sémantiques. Les lexies prédicatives sont celles dont le sens implique des participants; c'est le cas de tous les verbes (*X trébuche*, *X concède Y à Z*), mais aussi des adjectifs (*X est séduisant*), des adverbes (*X se déroule lentement*), des prépositions (*X sur Y*) et de plusieurs noms (*X gendre de Y*, *assaisonnement de X par Y*, *guerre entre X et Y*). Pour bien maîtriser le sens de ces lexies, il faut connaître les contraintes sémantiques qui pèsent sur ces participants. Par exemple, il faut que la lexie qui actualise *X* auprès de TRÉBUCHER dénote un être animé (*Ma grand-mère a trébuché*, mais pas **Ma statue a trébuché*). Il est donc important de spécifier ces contraintes dans la définition d'une lexie prédicative, ce que la LEC appelle le typage des actants. La définition de RAMPER ci-dessus en rend compte par 'en parlant d'animaux, de l'homme'.

Évidemment, ces différentes composantes de la définition pourront entraîner des erreurs lexicales de nature différente. De façon générale, nous pouvons affirmer que le sens est aussi soumis à la norme prescriptive, puisque les différents sens associés aux vocables sont décrits dans les dictionnaires et sont plutôt statiques. Par contre, les emplois figurés laissent une plus grande latitude au locuteur en ce qui concerne le sens.

Le cas le plus grave de transgression du sens est le non-respect du genre prochain, que nous appelons impropiété. Dans ce cas, la lexie utilisée ne véhicule pas du tout le sens visé par le scripteur. Dans certains cas, cette mauvaise association entre signifiant et signifié peut sembler difficilement explicable (ex. *frugal* pour exprimer 'abondant'), alors que dans d'autres cas, la confusion peut être de nature paronymique (*un départ *éminent*) ou encore découler de l'emploi d'anglicismes (*Elle est *rude* [= 'impolie'] *avec ses collègues*).

Dans d'autres cas, c'est plutôt la méconnaissance de certaines nuances sémantiques qui entraîne une erreur lexicale. Par exemple, l'utilisation douteuse de l'adjectif ROUTINIER dans *Il adorait ses échanges #routiniers avec sa grand-mère*. s'explique par le non-respect d'une différence spécifique. Devant ce cas, il y a lieu de croire que l'étudiant voulait plutôt exprimer le sens 'quotidien'; les deux sens sont assez proches, puisqu'ils qualifient tous deux quelque chose qui appartient à la vie de tous les jours, mais ROUTINIER véhicule en plus le sens de 'répétitif et monotone', ce qui ne convient pas réellement au contexte de la phrase.

Un troisième type d'erreur lexicale relatif au sens est le non-respect du typage des actants ; la confusion entre les verbes EMMENER et APPORTER en est un cas typique. Dans une phrase comme *Lucien a apporté sa petite sœur Ritha en camping.*, le typage du complément du verbe n'est pas respecté, puisque celui-ci doit dénoter un objet.

Un dernier type de problème sémantique mérite d'être traité, même s'il ne découle pas nécessairement d'une méconnaissance du sens d'une unité lexicale, mais naît plutôt de la cooccurrence de deux lexies. Il s'agit du pléonasme (*Ils cohabitent *ensemble depuis octobre.*), dans lequel un des deux termes reprend une composante de sens de l'autre.

3.2.3. Les propriétés de combinatoire

Polguère (2003: 37) définit les propriétés de combinatoire d'une unité lexicale comme « les contraintes propres à ce signe qui limitent sa capacité à se combiner avec d'autres signes linguistiques et qui ne peuvent être déduites ni de son signifié ni de son signifiant. » Ces propriétés s'articulent selon deux axes : la combinatoire grammaticale – c'est-à-dire les propriétés qui régissent le comportement morphosyntaxique de la lexie – et la combinatoire lexicale – les liens privilégiés qu'entretient l'unité lexicale avec les autres lexies.

Les propriétés de combinatoire grammaticale sont nombreuses et variées et, pour la plupart, elles sont soumises à la norme prescriptive et apparaissent dans les dictionnaires. Ce sont des informations que l'on doit connaître pour maîtriser pleinement une lexie et leur transgression entraîne une erreur. Des cas fréquents de ce type de propriétés sont la partie du discours (*Jessica est *alcoolisme.*), le genre nominal (**une écran d'ordinateur*), et l'invariabilité (**une obsèque émouvante*), mais il peut exister des propriétés plus rares, comme le fait pour une lexie de conditionner un mode verbal particulier (BIEN QUE + subjonctif) ou encore le fait de ne s'utiliser qu'en tournure négative (ex. SOURCILLER). Une autre propriété importante des unités lexicales est le régime, c'est-à-dire le fait d'avoir ou non des compléments, le caractère obligatoire de ces compléments et la présence ou l'absence de prépositions pour les introduire. La notion de régime verbal est bien connue, mais beaucoup de lexies appartenant à d'autres parties du discours possèdent aussi un régime (*X enclin à Y, acquiescement de X à Y*, etc.). Évidemment, le fait de ne pas respecter le régime entraînera une erreur (ex. *Ils faisaient partie *dans la même équipe.*).

Les erreurs découlant d'une violation d'une propriété de combinatoire grammaticale sont souvent considérées comme des « fautes de grammaire » par les enseignants. Comme la source des problèmes mentionnés plus haut est la méconnaissance de propriétés lexicales, nous croyons plus juste de parler d'erreurs lexicogrammaticales (ou lexicosyntaxiques dans le cas de problèmes de régime). Bien que ces problèmes se manifestent syntaxiquement ou morphologiquement, aucune règle générale de la grammaire du français ne permettra de les expliquer et il faudra nécessairement avoir recours à une propriété lexicale pour y parvenir. Par exemple, nous jugeons beaucoup plus pertinent, devant le syntagme **une violente ouragan*, de mentionner à l'étudiant que le nom OURAGAN est masculin que de considérer qu'il ne maîtrise pas les règles d'accord du déterminant et de l'adjectif avec le nom.

Le deuxième axe de combinatoire est la combinatoire lexicale, qui se manifeste principalement par la collocation, phénomène très fréquent dans les langues. Il s'agit de la cooccurrence de deux unités lexicales selon le patron suivant : une lexie, appelée base de la collocation, est choisie librement pour le sens qu'elle véhicule, alors que l'autre, le collocatif, est choisie en fonction de la base pour exprimer un sens particulier auprès d'elle (voir Hausmann (1979)). Par exemple, dans *faire un choix*, la base CHOIX est choisie librement alors que le collocatif FAIRE est en quelque sorte imposé par ce choix (!). En effet, le synonyme DÉCISION aurait demandé un collocatif différent : *prendre une décision* et non **faire une décision*. La collocation est un phénomène arbitraire, c'est-à-dire qu'on ne peut anticiper les collocatifs d'une lexie particulière; c'est l'usage qui « fixe » les collocations.

Par conséquent, il faut connaître les collocatifs d'une lexie que l'on désire utiliser, tâche d'autant plus difficile que ce type d'information n'apparaît pas systématiquement dans les dictionnaires. En effet, seule la consultation de corpus permettra de statuer sur l'acceptabilité d'une collocation; cette dimension est donc soumise à la norme descriptive. Évidemment, l'emploi de mauvais collocatifs est perçu comme problématique, comme dans les exemples suivants : *Elle leur *interroge plusieurs questions. Un film ?énormément drôle!*

3.2.4. D'autres types de problèmes lexicaux

Nous avons vu, au point 3.1, que certaines contraintes pragmatiques pouvaient aussi occasionner des erreurs lexicales en vertu de la norme fonctionnelle. Par exemple, le registre de langue employé doit être en adéquation avec la situation de communication. Il est aussi important qu'il y ait une certaine homogénéité dans le registre employé dans un texte et une cassure par rapport à ce registre sera perçue comme maladroite.

De plus, certains contextes d'écriture demanderont l'emploi d'un français international, c'est-à-dire exempt de termes propres à une variété de français particulière qui pourraient constituer une entrave à la compréhension pour des lecteurs qui lui sont étrangers. Utiliser des termes propres au français québécois standard, comme DÉBARBOUILLETTE ou ACHALANDAGE, pourrait poser problème si le texte s'adresse à un lectorat international.

Finalement, l'emploi d'un vocabulaire pauvre, imprécis ou peu varié peut aussi être considéré comme un problème de nature lexicale, puisque cela témoigne d'un manque de vocabulaire. Par contre, les critères pour juger de tels problèmes sont plutôt flous et l'appréciation de la qualité ou la richesse du vocabulaire est souvent très subjective.

4. Description des problèmes lexicaux d'étudiants

Une fois notre typologie des problèmes lexicaux proposée, nous avons réalisé notre deuxième objectif, qui visait à décrire les problèmes rencontrés dans les rédactions d'étudiants universitaires. Cette étape a en quelque sorte servi à valider notre typologie en démontrant son utilité comme outil d'aide au diagnostic.

4.1. Méthodologie

Pour cette deuxième étape de la recherche, nous avons dépouillé un corpus de 103 textes d'étudiants et en avons extrait tous les problèmes lexicaux. L'ensemble des textes ont été rédigés, dans le cadre d'un cours de français écrit pour futurs enseignants, par des étudiants inscrits en première année de baccalauréat¹ en éducation et qui ont eu le français comme seule langue de scolarisation depuis le primaire. La rédaction d'une page environ portait sur un sujet imposé : résumé d'un film, description d'un endroit ou texte sur la discipline dans les écoles. Le travail a été réalisé à la maison – le recours aux ouvrages de référence était donc permis – et remis au professeur en tant qu'évaluation formative. Tous les problèmes lexicaux ont été codés à l'aide de la typologie présentée plus haut par l'auteur de la présente recherche. Un peu plus du quart des problèmes (111) a ensuite fait l'objet d'un contre-codage par une collègue. Le taux d'accord interjuges était de 86,2 %, ce qui témoigne du caractère opérationnel de la typologie.

4.2. Résultats et discussion

L'analyse du corpus a révélé un total de 391 problèmes lexicaux – c'est-à-dire environ 3,8 problèmes par 250 mots – et plus de 99 % des cas relevés ont pu être associés à l'une ou l'autre des classes d'erreurs de notre typologie.

¹ Études de premier cycle universitaire au Québec.

Les problèmes relatifs à la combinatoire sont les plus fréquents (30,7 %). Parmi ceux-ci, 23,5 % concernent la combinatoire grammaticale; il s'agit surtout de problèmes de régime – notamment avec les verbes (**elle lui informe ses déplacements*) – et du non-respect de l'invariabilité (**Vous êtes amateur d'actions ?*) ou du genre nominal (**un idole*). De telles informations apparaissent pourtant dans les dictionnaires et ces erreurs auraient pu être évitées en les consultant. Les problèmes de combinatoire lexicale (collocation), quant à eux, représentent 7,2 % du total; les collocations les plus problématiques semblent être celles qui ont un nom pour base et un verbe comme collocatif (**Les élèves **appliquent** peu de respect envers les enseignants.*). Le désir d'éviter d'utiliser des verbes qui sont pourtant des collocatifs fréquents (ex. FAIRE, AVOIR), mais que les enseignants qualifient souvent de verbes « pauvres », pourrait peut-être expliquer ces erreurs.

Les problèmes de sens, avec 30,4 % du total, sont aussi très fréquents. Bien que notre typologie permette de distinguer théoriquement des problèmes de sens distincts (voir 3.2.2.), nous constatons qu'il est difficile de les départager lors de l'analyse de cas réels. En effet, le sens étant assez abstrait, il n'est pas toujours facile de pointer la composante de sens à l'origine d'une utilisation douteuse. De plus, une seule occurrence d'utilisation ne témoigne pas de façon très détaillée de la conception du sens chez le scripteur. Des recherches plus poussées sur le diagnostic de problèmes de sens permettraient peut-être d'en développer une méthode d'analyse plus efficace. Néanmoins, il est intéressant de trouver un si grand nombre de problèmes sémantiques chez des étudiants dont le français est la langue maternelle.

Sans grande surprise, nous constatons aussi la présence de nombreux problèmes relatifs à la forme (21 %), dont l'écrasante majorité sont des erreurs d'orthographe lexicale. Bien que peu surprenante, cette constatation est tout de même préoccupante si l'on considère que les textes ont été rédigés à la maison et que le recours aux ouvrages de référence était permis. Ceci témoigne peut-être d'une certaine négligence de la part des étudiants, qui se disent sans doute qu'une « simple » faute d'orthographe n'entrave pas la compréhension.

Les problèmes de registre de langue représentent 9 % des cas et se manifestent la plupart du temps par l'emploi de termes familiers (CRAMER, À CAUSE QUE, *elle angoisse*). Pourtant, le dictionnaire note l'appartenance à des registres de langue particuliers et indique le terme standard correspondant. Par contre, il se peut que le sujet du texte ait eu une influence sur le ton plus familier de certaines rédactions.

Les problèmes ayant trait à la richesse et à la variété du vocabulaire représentent 5,6 % du total et prennent généralement la forme de répétition. Quant à la richesse du vocabulaire, très peu de cas ont été relevés, faute de méthode claire pour statuer sur ce type de problèmes.

Deux pour cent des erreurs relevées pouvaient donner lieu à un double codage, c'est-à-dire qu'elles appartenaient à plus d'un type. En effet, certains problèmes pouvaient s'expliquer de différentes façons et l'impossibilité d'avoir accès aux pensées du scripteur nous empêchait d'opter pour une interprétation. Ceci ne découle pas selon nous de problèmes avec notre typologie, mais témoigne plutôt de la difficulté d'expliquer une erreur sans avoir accès aux processus qui ont conduit à la produire. Finalement, 1,3 % des problèmes relevés ont été considérés plutôt comme des coquilles attribuables à des erreurs d'inattention.

Conclusion

Différentes conclusions émanent de notre travail. D'un point de vue théorique, nous croyons avoir démontré qu'une délimitation précise de la notion d'erreur lexicale est possible et que la TST offre un cadre théorique favorable à cet objectif. Malgré le caractère opérationnel de la typologie des problèmes lexicaux proposée, nous constatons par contre que certains types de problèmes, notamment ceux liés au sens, demeurent difficiles à diagnostiquer.

D'un point de vue plus didactique, l'analyse des problèmes lexicaux des étudiants offre certaines pistes. Premièrement, le nombre important d'erreurs concernant les propriétés de combinatoire devrait nous convaincre de la nécessité d'accorder plus d'importance à cette dimension des unités lexicales dans l'enseignement, particulièrement aux notions de régime et de collocation, qui sont en général peu abordées en classe. La présence élevée de problèmes de sens nous pousse aussi à nous interroger sur la capacité des étudiants à remettre en question leur compréhension du sens des lexies qu'ils utilisent. Leur habileté à bien différencier les registres de langue et à adapter leur écrit au contexte scolaire est aussi remise en question. Une bonne initiation à l'utilisation de dictionnaires, par la présentation de leur structure et de la variété des informations présentées, devrait sensibiliser les étudiants à l'étendue des connaissances en jeu dans l'utilisation du vocabulaire.

Évidemment, la typologie que nous proposons ne prétend pas être un outil infallible qui permettrait de diagnostiquer aisément tout problème lexical. En revanche, nous croyons que notre travail met en lumière l'importance de s'intéresser à ce type de problèmes somme toute assez présents dans les rédactions d'étudiants et très peu étudiés jusqu'à ce jour.

Références

- Andidote Prisme*, version 6 [logiciel], Montréal, Druide informatique, 2006.
- CATACH, Nina, GRUAZ, Claude & DUPREZ, Daniel (1986). *L'orthographe française : traité théorique et pratique avec des travaux d'application et leurs corrigés*. Paris: Nathan.
- DALLAIRE, Édouard (1997). *Évolution du discours institutionnel tenu sur la langue, la norme et la grammaire depuis le Rapport Parent (Mémoire de maîtrise du 2e cycle non publié)*, Université du Québec à Montréal, Montréal.
- DENANCE, Michel (1999). *L'évaluation des fautes de vocabulaire à l'évaluation de l'épreuve uniforme*, Correspondance, 4, 4.
- ENGBER, Cheryl A. (1995). The Relationship of Lexical Proficiency to the Quality of ESL Compositions, *Journal of Second Language Writing*, 4, 2: 139-155.
- HAUSMANN, Franz Josef (1979). Un dictionnaire des collocations est-il possible ?, *Travaux de littérature et de linguistique de l'Université de Strasbourg*, XVII, 1: 187-195.
- LEFRANÇOIS, Pascale (2004). Les stratégies de résolution de problèmes orthographiques d'étudiants à l'université, in *La littératie : conceptions théoriques et pratiques de l'enseignement de la lecture-écriture*, Barré-De Miniac, Christine, Brissaud, Catherine & Rispaill, Marielle (Eds), Paris: l'Harmattan, 233-244.
- LEGENDRE, Renald (2005). *Dictionnaire actuel de l'éducation* (3e ed.). Montréal: Guérin.
- LÉPINE, François (1995). Bilan des tests de français à l'admission aux universités québécoises (1987-1994), *Revue des sciences de l'éducation*, 21, 1: 17-36.
- MEL'ČUK, Igor, CLAS, André & POLGUÈRE, Alain (1995). *Introduction à la lexicologie explicative et combinatoire*. Louvain-la-Neuve: Duculot.
- MEL'ČUK, Igor (1997). *Vers une linguistique sens-texte : leçon inaugurale faite le vendredi 10 janvier 1997*. Paris, Collège de France: Chaire internationale.
- PATRI, Jacques & VANSNICK, Nicole (1992). Comment aborder les fautes lexicales en classe, *Enjeux*, 26: 46-55.
- POLGUÈRE, Alain (2003). *Lexicologie et sémantique lexicale : notions fondamentales*. Montréal: Presses de l'Université de Montréal, Coll. Paramètres.
- POLGUÈRE, Alain (2004). *Savoir consulter un dictionnaire, c'est bien ; savoir comment le construire, c'est mieux*, *Québec français*, 134: 68-70.
- ROY, Gérard-Raymond, LAFONTAINE, Louise & LEGROS, Catherine (1995). *Le Savoir grammatical après treize ans de formation*. Sherbrooke: Éditions du CRP.
- SAUSSURE, Ferdinand (de) (1960). *Cours de linguistique générale*. Paris: Payot.

METTRE EN ŒUVRE SES COMPÉTENCES DISCURSIVES EN L2 : QUELLES OPPORTUNITÉS EN MILIEU SCOLAIRE ?

Évelyne BERGER
evelyne.berger@unine.ch
 Centre de Linguistique Appliquée
 Université de Neuchâtel
 Espace Louis-Agassiz 1
 2000 Neuchâtel
 Suisse

Abstract : This study focuses on turn-taking practices in various oral activities in the French second language classroom. It examines whether French second language learners have the opportunity to spontaneously take the floor in the language classroom. It attempts to identify the interactional conditions that are potentially favourable for the development of discourse-organisational skills.

1. Introduction

La participation aux interactions verbales suppose la capacité à prendre la parole au moment opportun et de manière appropriée. Cette capacité à enchaîner sur le discours d'autrui fait partie d'un ensemble complexe et diversifié de compétences discursives permettant aux locuteurs de gérer conjointement le discours au cours de son déroulement. Pour des apprenants de langue seconde, la prise de parole est souvent problématique. Elle constitue par conséquent un obstacle à la communication. Dans cette perspective, l'analyse des prises de parole en classe de langue peut nous renseigner sur les conditions interactives que diverses situations didactiques offrent pour la mobilisation (et le développement éventuel) des compétences discursives.

La présente étude se propose d'analyser les pratiques de gestion des tours de parole dans des activités didactiques variées au sein de la classe de français langue seconde. Nous nous intéresserons en particulier aux enchaînements rapides. Les ressources linguistiques et interactives nécessaires à leur réalisation en font des indicateurs de conditions propices à la mobilisation des compétences discursives. Après avoir présenté une définition des compétences discursives et la problématique liée à leur mise en œuvre en langue seconde (§2), nous mènerons une analyse tant quantitative que qualitative des contextes d'occurrence des enchaînements rapides (§3). Nous terminerons par des considérations sur les retombées didactiques des résultats esquissés (§4).

2. Les compétences discursives dans l'interaction¹

2.1. Définir les compétences discursives

Dans la communication en face-à-face, les interactants sont amenés à défendre leur point de vue face à autrui, argumenter, résoudre des malentendus, gérer des désaccords, s'ajuster mutuellement, etc. Ces activités socio-interactionnelles impliquent un engagement de la part

¹ Cette étude s'inscrit dans un projet de recherche, mené au Centre de linguistique appliquée de l'Université de Neuchâtel (Suisse) et dirigé par S. Pekarek Doehler, qui s'intitule: «L'organisation du discours-dans-l'interaction en langue première et seconde : acquisition, enseignement, évaluation » (dans le cadre du PNR 56; FNS 405640-108663/1).

des interactants dans la gestion et la construction conjointe du discours. Cet engagement présuppose des compétences discursives qui reposent sur la mobilisation appropriée de ressources interactives et linguistiques dans la dynamique de l'interaction (Pekarek, 1999) et en cela une capacité d'adaptation au cours des événements communicatifs.

La communication ne consiste pas uniquement en la mise en œuvre de connaissances linguistiques, mais comprend également une dimension stratégique et sociolinguistique, qui, à son tour, est composée de connaissances socioculturelles et discursives, comme l'ont suggéré Canale et Swain (1980) dans leur définition de la compétence de communication. Pour ce qui est de la dimension discursive, Canale et Swain (1980) se réfèrent à la cohérence entre les énoncés, mais uniquement sur un plan monologique. Dans la présente étude, nous opterons pour une approche dialogique des compétences discursives inspirée notamment des travaux issus de l'analyse conversationnelle d'origine ethnométhodologique.

Les compétences discursives, dans cette optique, interviennent à différents niveaux du discours : au niveau local (gestion des thèmes conversationnels et des référents, enchaînements entre les tours de parole, etc.) et au niveau global (gestion des ouvertures et clôtures de la conversation ou d'épisodes conversationnels). L'une de ces compétences est notamment la capacité à enchaîner sur le discours d'autrui, laquelle implique une coordination minutieuse des échanges verbaux s'inscrivant dans un système d'alternance des locuteurs (Sacks *et al.*, 1974). Les participants se basent sur des éléments de nature syntaxique, pragmatique et prosodique pour coordonner leurs activités discursives dans l'interaction (Sacks *et al.*, 1974 ; Schegloff, 1996 ; Ford & Thompson, 1996).

2.2. Acquérir les compétences discursives

Pour les apprenants d'une langue seconde, l'acquisition de ces compétences discursives semble être problématique. En effet, différentes études récentes issues du fonctionnalisme discursif ont décrit l'organisation du discours comme une zone de résistance dans l'apprentissage (Carroll *et al.*, 2000 ; Carroll & Lambert, 2003 ; Murcia-Serra, 2003). Elles montrent que les apprenants avancés d'une langue seconde maîtrisent parfaitement les formes linguistiques dans leur adéquation formelle mais ont des difficultés à utiliser ces structures linguistiques de manière appropriée dans des tâches communicatives complexes. En fait, le choix des formes linguistiques au sein d'un ensemble de formes disponibles et leur adéquation au contexte repose souvent sur un schéma provenant de la langue première. Ce constat implique au niveau développemental que l'apprenant d'une langue seconde n'intègre les schémas d'organisation du discours de la langue seconde que tardivement. Pour y parvenir, l'apprenant doit nécessairement passer par une étape de reconceptualisation des principes organisationnels liés à sa langue première (Carroll *et al.*, 2000). Ainsi, ces études montrent de manière évidente que l'acquisition d'une langue seconde ne consiste pas simplement en l'apprentissage de formes linguistiques mais que dans la pratique effective de la langue seconde, ces formes doivent être utilisées de manière appropriée au contexte et aux modes d'organisation du discours.

Alors que les phénomènes susmentionnés n'ont jusqu'ici été étudiés que dans les productions narratives, d'autres observations émanant de travaux sur les échanges interactifs en langue seconde confirment le caractère problématique des compétences discursives comme objet d'acquisition. En effet, malgré une compétence linguistique développée au niveau lexical et syntaxique, les apprenants avancés d'une langue seconde ont souvent beaucoup de difficultés à prendre une part active dans les interactions où les tours de parole s'enchaînent rapidement et de manière peu prévisible (Pekarek, 1999).

Dans les interactions en classe de langue, les apprenants ont souvent des difficultés à prendre la parole spontanément. Pourtant, s'approprier le tour de parole dans l'interaction en cours est la condition *sine qua non* pour, à la fois, mettre en pratique ses compétences, et donc les

rendre visibles en classe, et pour éventuellement les développer au travers de l'activité discursive (Pekarek Doehler, 2005). Or, la difficulté pour les apprenants d'une langue seconde réside dans le fait que la capacité à enchaîner sur autrui mobilise des ressources interactives communes à la langue première et à la langue seconde, mais aussi des ressources linguistiques spécifiques à la langue seconde. La difficulté peut se situer tant au niveau du repérage de la place pertinente pour le changement de locuteur qu'au niveau du choix des marques linguistiques pour articuler son tour de parole au contexte discursif. D'autre part, on suppose que cette capacité à enchaîner sur le discours d'autrui a été acquise au travers de la langue première et qu'elle est déjà maîtrisée par les apprenants de la langue seconde. Mais enchaîner sur le discours d'autrui peut devenir problématique lorsque les ressources linguistiques nécessaires pour y arriver sont encore en voie de développement. Enfin, la participation à un échange verbal présuppose également une certaine fluidité verbale de la part des interactants, ce qui pose souvent problème chez les apprenants de langue seconde.

Dans cette perspective, l'étude des pratiques d'enchaînement des tours de parole est particulièrement intéressante dans la mesure où elle a un effet de loupe sur le travail de coordination et d'ajustement des activités des locuteurs. Elle apporte donc un éclairage sur la manière dont les apprenants accomplissent ce travail interactif alors qu'ils ont des moyens linguistiques plus ou moins lacunaires à différents stades de développement de la langue seconde.

3. La gestion des tours de parole sous la loupe

3.1. Prendre la parole : une coordination minutieuse

La coordination des échanges s'inscrit dans un système d'alternance des locuteurs qui a été décrit par Sacks, Schegloff et Jefferson (1974). Typiquement, la transition entre les tours de parole est réalisée de manière fluide sans chevauchement ou long silence mais avec une pause minimale. Toutefois, il arrive que les interactants s'écartent de ce principe : le changement de locuteur n'a pas lieu à des points de transition pertinents (*transition relevant places*) et peut constituer une «violation» du système d'alternance des locuteurs (Ford & Thompson, 1996 : 159). Ces déviations reflètent une manière alternative d'utiliser le système d'alternance des locuteurs.

Il arrive également que les locuteurs enchaînent sur le tour de parole d'un locuteur précédent sans qu'il y ait une pause minimale mais toutefois sans chevaucher le tour précédent. Il s'agit dans ce cas d'une transition sans césure que l'on appellera *enchaînement rapide*. La réalisation de ces enchaînements rapides repose sur la mobilisation de ressources linguistiques et interactives. En effet, les locuteurs doivent d'abord identifier de manière précise le point de transition pertinent pour le changement de locuteur. Ce point représente le moment où une unité de construction de tour (*turn constructional unit*, désormais *TCU*) est complète dans le contexte discursif du point de vue syntaxique, sémantico-pragmatique et prosodique (Ford & Thompson, 1996). Selon Schegloff (1987), la forme et le type du *TCU* peuvent être projetés dès le début du tour par le choix des unités linguistiques, ce qui permet aux interlocuteurs d'en anticiper la fin. Pour ce faire, les locuteurs s'orientent vers les aspects syntaxiques et prosodiques de la forme que prend le tour au cours de sa réalisation ainsi que sa complétude sémantico-pragmatique (par exemple, un simple «oui» placé après une question de type «oui-non» peut être considéré comme une unité pragmatiquement complète dans ce contexte séquentiel spécifique). L'enchaînement rapide implique que le prochain locuteur enchaîne sur le locuteur en cours dès que le *TCU* arrive à son point de complétude, empêchant ainsi tout autre candidat potentiel de prendre la parole ou de poursuivre sur un nouveau *TCU*. En fait, les locuteurs placent avec précision leur tour de parole dans le

déroulement séquentiel du discours (Jefferson, 1973), comme nous pouvons le constater dans l'exemple suivant² :

Extrait 1 : L2Secl-ts-tg1-gs

1 Enq: ce soir vous dites euh on est invité à une fête et
maintenant
2 il faut qu'on discute ça=
3 > Ani: =mais c'est difficile

Dans l'extrait ci-dessus, Anila (Ani) ne prend pas la parole en chevauchant le locuteur en cours mais précisément au moment où le tour de parole est complet sur les plans syntaxique et pragmatique. Le tour de parole de l'enquêtrice (Enq) étant construit en deux parties articulées par le marqueur de coordination *et*, Anila n'enchaîne qu'à la fin de la deuxième partie, qui ne constitue pas seulement le point de complétude d'un syntagme verbal, mais aussi d'une construction syntaxique complexe. De plus, du point de vue pragmatique, l'exemple que l'enquêtrice donne aux apprenantes pour comprendre l'activité est également terminé au moment où Anila prend la parole. Enfin, l'enchaînement est effectué immédiatement, c'est-à-dire sans la moindre césure entre les deux tours de parole, ce qui est indiqué par le symbole typographique « = ».

Dans la mesure où les enchaînements rapides sont accomplis avec une précision remarquable, ils présupposent nécessairement une capacité de coordination des échanges verbaux, qui, de plus, est réalisée de manière fluide. Dans le cas des interactions en langue seconde, ils peuvent être un indice de conditions interactives potentiellement riches en occasions pour le déploiement des compétences discursives en général. C'est pour cette raison que la présente recherche se propose d'étudier les enchaînements rapides produits par des apprenants d'une langue seconde dans différentes activités à but communicatif en classe de langue, à deux moments de l'apprentissage.

3.2. Les données

Le corpus sur lequel porte la présente étude exploratoire est constitué de quatre heures d'enregistrements audio-vidéo. Les données concernent des apprenants de français L2 vivant en région germanophone de Suisse. Ils se répartissent en deux groupes : un niveau moyen (école obligatoire) où les apprenants, âgés de 13-14 ans, suivent un enseignement de français L2 depuis 4 ans, 4-5h par semaine, et un niveau avancé (lycée) où les apprenants, âgés de 16-18 ans, suivent un enseignement de français L2 depuis 6-8 ans, 3-5h par semaine. Les interactions ont été observées dans deux types d'activités : le premier type (exercice de grammaire à l'oral, discussion de littérature) regroupe des discussions qui sont fortement régulées par l'enseignant (au niveau de l'attribution des tours de parole et de la gestion des contenus thématiques) impliquant une certaine prévisibilité dans l'organisation des échanges. Le deuxième ensemble d'activités (travail en groupe, débat) regroupe des discussions plus libres où les apprenants participent à la gestion et à la construction du discours. Le corpus étudié est ainsi composé d'une leçon par type d'activités et par groupe d'apprenants.

4. Les enchaînements rapides : traces d'un engagement discursif

4.1. Activités didactiques et formats d'interaction

Une première analyse quantitative (cf. tableau ci-contre) montre que le nombre d'enchaînements rapides est nettement plus élevé dans le travail en groupe pour le niveau moyen et dans le débat pour le niveau avancé. Ce résultat suggère que la fréquence de ce

² Cf. conventions de transcription en fin d'article

phénomène discursif est étroitement liée au format d'interaction. Ainsi, ce phénomène augmente lorsque la responsabilité de la gestion de l'interaction (organisation des échanges et des contenus thématiques) est partagée par les apprenants, comme c'est le cas dans le travail en groupe ainsi que le débat. Cette prise de responsabilité dans l'attribution des tours de parole notamment est manifestée par le fait qu'ils se sélectionnent spontanément à la première occasion (dès la fin d'un tour de parole) comme prochains locuteurs sans attendre d'être désignés par l'enseignant comme c'est le cas habituellement dans les interactions en classe. Dans cette, perspective, la quantité d'enchaînements rapides devient un indice d'un engagement discursif de leur part dans l'interaction.

Activités	Niveau moyen		Niveau avancé	
	Discussion gérée par le prof. (ex. gramm.)	Discussion libre (trav. en groupe)	Discussion gérée par le prof. (disc. de litt.)	Discussion libre (débat)
Nb ER	15	30	3	8

Résultats quantitatifs - Nombre d'enchaînements rapides (ER) dans une leçon par groupe d'apprenants et par type d'activité

Ces observations rejoignent les résultats de Pekarek (1999) qui montre que les différentes modalités de collaboration dans la gestion du discours en classe avancée de français L2 ont des retombées importantes sur la nature du travail discursif engagé : les activités didactiques qui mettent en place un format d'interaction basé sur une gestion mutuelle de l'organisation des échanges et qui se centrent sur l'échange de contenus non prévisibles invitent les apprenants à accomplir un travail discursif plus complexe.

4.2. Enchaînements rapides et organisation socio-interactionnelle

Dans la mesure où les enchaînements rapides constituent une exception quant aux méthodes d'alternance des locuteurs (le cas «typique» étant un enchaînement avec une mini-césure), on peut supposer que leur occurrence résulte du fait que le locuteur poursuit des enjeux communicatifs spécifiques (Ford & Thompson, 1996 :164). Par conséquent, leur signification interactive n'est interprétable que localement, raison pour laquelle une analyse de leur placement dans le déroulement séquentiel de l'interaction est nécessaire pour comprendre leur rôle dans l'organisation socio-interactionnelle de la conversation. Prenons l'extrait suivant:

Extrait 2: L2SecI-ts-tg1-gs

1 Ebr: ähm:: (.) kömm mr ufschribe?
2 Enq: oui vous pouvez prendre des notes
3 Ebr: oui=
4 Enq: =parce que après vous allez présenter (.) vous
choisissez un
5 des deux
6 Ebr: mhm
7 Enq: ce soir vous dites euh on est invité à une fête et
maintenant
8 il faut qu'on discute ça=
9 > Ani: =mais c'est difficile
10 Enq: ça je sais (.) mais je sais aussi que vous êtes capables
de
11 faire pleins de choses et on veut voir ce que vous êtes
12 capables de faire

Dans cet extrait, les apprenants (qui sont de niveau moyen) ont dû se mettre en groupe de trois ou quatre pour la réalisation d'une activité de groupe. Suite à la question posée par Ebru (Ebr) à la ligne 1 en suisse-allemand sur le fait de pouvoir prendre des notes écrites pendant l'activité, l'enquêtrice explique la manière de procéder à l'exercice, en la justifiant (lignes 4-5) et en donnant un exemple (lignes 7-8). Or, à la ligne 9, Anila conteste l'activité proposée,

à savoir discuter en français d'un thème donné. Son positionnement contrastif est introduit par le marqueur discursif *mais* et suit immédiatement le propos contesté dans le tour de parole précédent au moyen d'un enchaînement rapide.

Chez les apprenants plus avancés, on observe également la présence d'enchaînements rapides lors de positionnements contrastifs, comme l'illustre l'exemple suivant:

Extrait 3: L2SecII-ob4cIII

1 Pro: +vous vouliez dire encore quelque chose((très
doucement))+
2 Ben: je voulais juste voulu: euh: dire à peu près la même
chose
3 (...) je crois que: ça n'existe pas un pessimisme (...)
euh
4 un pessimiste absolu (..) je crois tout le monde a un
peu:
5 d'optimisme et:=
6 > Ale: =mais peut-on seulement discuter sur quelque chose qui
exi-
7 euh qui existe.
8 (2.8)
9 Ben: euh: tu (.) vous pouvez (.) tu peux discuter (...)
beaucoup
10 de choses qui ne: n'existent pas.

Aux lignes 6-7, Alex (Ale) manifeste son désaccord avec son collègue Benjamin (Ben) en contre-argumentant son propos. Sa prise de parole contrastive est réalisée au moyen d'un enchaînement rapide et de *mais* en début de tour, lequel permet ainsi d'opposer le tour de parole en cours à celui qui précède. On observe ainsi la présence des mêmes éléments que dans l'extrait précédent pour la réalisation d'un enchaînement contrastif.

En fait, les enchaînements rapides sont souvent réalisés dans des séquences de désaccord. Il est dès lors possible de penser que dans ces moments-là, l'apprenant est motivé par un enjeu affectif important qui le pousse à s'engager plus intensément dans la construction et la gestion du discours. Il puise alors au maximum dans les ressources linguistiques et interactives qu'il a à sa disposition pour faire valoir son point de vue, mettant ainsi en œuvre ses compétences discursives. Dans ce dernier extrait, Alex s'approprie la parole alors que le locuteur en cours n'a pas terminé. En effet, l'enchaînement rapide est réalisé à un point de transition possible (complétude potentielle du *TCU* en cours) mais où le locuteur en cours voulait continuer comme le signale la présence du *et allongé* qui indique une suite à venir (ligne 5). Ainsi, le fait que Alex saisisse la première occasion pour manifester son désaccord, au détriment du locuteur en cours, témoigne de son engagement dans le discours.

D'autre part, les enchaînements rapides peuvent également servir de stratégie interactive pour s'approprier la parole dans des moments de compétition pour la parole, comme on peut le voir dans les extraits cités. Dans l'extrait 2, le tour de parole qui comprend un enchaînement rapide est produit par Anila, laquelle se trouvait en marge de l'interaction à ce moment-là. En effet, la discussion semble se dérouler entre l'enquêtrice et Ebru, laquelle lui avait posé une question. Le fait qu' Anila prenne la parole dès la fin du tour de parole de l'enquêtrice est un moyen pour réintégrer la discussion en se réappropriant le statut de locutrice. Dans l'extrait 3, les interactants se trouvent également dans une situation de compétition pour la parole qui est très typique des débats, étant donné que cette activité implique de faire valoir son point de vue par rapport aux autres : l'enchaînement rapide réalisé par Alex semble servir de moyen pour s'approprier la parole alors qu'il n'avait pas été invité à intervenir.

5. Implications didactiques

Comme nous avons pu l'observer précédemment, les enchaînements rapides sont une manifestation visible d'une fluidité dans la coordination des échanges verbaux. La précision séquentielle de l'enchaînement est essentielle pour que le changement de locuteur puisse se faire sans interruption. Elle est rendue observable à la surface de la conversation du fait qu'elle est accomplie plus tôt que les changements réguliers, donc sans césure inter-tours minimale. De plus, cette réalisation particulière du changement de locuteur est exploitée par les interactants comme ressource pour l'organisation socio-interactionnelle de la conversation : par exemple pour manifester son désaccord ou pour prendre la parole lorsqu'il y a compétition autour de la place de prochain locuteur.

Ainsi, la présence d'enchaînements rapides peut être le signe d'un engagement discursif des interactants dans la gestion et la construction conjointe du discours. Cet engagement discursif peut être jugé comme une condition favorable au développement des compétences discursives. Dans cette optique, l'observation de ce phénomène dans une variété d'activités didactiques et communicatives nous renseigne sur les activités communicatives potentiellement riches en occasions pour la mise en œuvre des compétences discursives. Plus les apprenants ont la possibilité de prendre en charge l'interaction, comme c'est le cas pour le travail en groupe et le débat, plus les enchaînements rapides sont fréquents. Ainsi, face à l'imprévisibilité du déroulement des échanges, les apprenants sont contraints de puiser dans toutes les ressources qu'ils ont à disposition pour adapter leur discours au cours des événements communicatifs.

Dans une perspective interactionniste de l'acquisition des langues, l'interaction n'est pas seulement un lieu de mise en pratique des connaissances et des savoir-faire en langue seconde mais est conçue «comme le lieu de leur émergence, de leur transformation et de leur remaniement» (Mondada et Pekarek, 2001 : 111). En effet, l'interaction sociale a un effet structurant sur le développement des connaissances et savoir-faire linguistiques du fait que l'apprenant est amené à accomplir différentes tâches et activités (Mondada & Pekarek, 2001) et à surmonter des obstacles à la communication (Krafft & Dausendschön-Gay, 1994). Ainsi dans notre cas, les interactions en classe sont un site essentiel de développement des compétences linguistiques des apprenants scolaires d'une langue seconde. Comme nous l'avons déjà remarqué, la capacité à enchaîner sur le discours d'autrui est essentielle dans toute communication en face-à-face, sans quoi l'apprenant ne peut pas prendre part activement à la conversation. Ainsi, dans l'éventualité où les activités didactiques ne donneraient pas les occasions de mettre en œuvre, et par là même de développer, cette capacité à enchaîner sur le discours d'autrui, les apprenants seraient alors privés de possibilités de manifester leur compétence en langue seconde dans le contexte scolaire. Dans cette optique, il est nécessaire de promouvoir les activités didactiques qui privilégient un partage de la gestion de l'interaction, tels que le travail en groupe ou le débat.

Conventions de transcription

Enq	Enquêtrice
Pro	Professeur
Ani	Anila (élève)
Ebr	Ebru (élève)
Ben	Benjamin (élève)
Ale	Alex (élève)
=	enchaînement rapide
(.)	pause très courte (moins de 0.3 seconde)
(..)	pause moyenne (0.3 à 0.6 seconde)
(...)	pause longue (0.6 à 0.9 seconde)
(2.8)	pause prolongée de 2.8 secondes
?	intonation montante
.	intonation descendante

: allongement d'une syllabe
 (()) commentaire de la part du transcripteur
 + début et fin de la séquence sur laquelle porte le commentaire du transcripteur
 > ligne dans laquelle se trouve un enchaînement rapide

Références bibliographiques

- CANALE, Michael & SWAIN, Merrill (1980). Theoretical bases of communicative approaches to second language teaching and testing, *Applied Linguistics*, 1: 1-47.
- CARROLL, Mary, MURCIA-SERRA, Jorge, WATOREK, Marzena & BENDISCIOLI, Alessandra (2000). The relevance of information organization to second language acquisition studies: the descriptive discourse of advanced adult learners of German, *Studies in second language acquisition*, 22: 441-466.
- CARROLL, Mary & LAMBERT, Monique (2003). Information structure in narratives and the role of grammaticised knowledge. A study of adult French and German learners of English, in *Information structure and the dynamics of language acquisition*, Dimroth, Christine & Starren, Marianne (Eds). Amsterdam/ Philadelphia: John Benjamins, 267-286.
- FORD, Cecilia E. & THOMPSON, Sandra A. (1996). Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns, in *Interaction and grammar*, Ochs, Elinor *et al.* (Eds). Cambridge : Cambridge University Press, 134-184.
- JEFFERSON, Gail (1973). A case of precision timing in ordinary conversation: overlapped tag-positioned address terms in closing sequences, *Semiotica*, 9, 1: 47-96.
- KRAFFT, Ulrich & DAUSENDSCHÖN-GAY, Ulrich (1994). Analyse conversationnelle et recherche sur l'acquisition, *Bulletin VALS-ASLA*, 59: 127-158.
- MONDADA, Lorenza & PEKAREK DOEHLER, Simona (2001). Interactions acquisitionnelle en contexte, *Le français dans le monde, numéro spécial juillet 2001*: 107-137.
- MURCIA-SERRA, Jorge (2003). Acquiring the linkage between syntactic, semantic and informational roles in narratives by Spanish learners of German, in *Information structure and the dynamics of language acquisition*, Dimroth, Christine & Starren, Marianne (Eds). Amsterdam/Philadelphia: John Benjamins, 289-310.
- PEKAREK, Simona (1999). *Leçons de conversation*. Fribourg : Editions Universitaires.
- PEKAREK DOEHLER, Simona (2005). De la nature située des compétences en langue, in *Repenser l'enseignement des langues : comment identifier et exploiter les compétences ?* Pouliot, Michèle *et al.* (Eds). Villeneuve d'Ascq : Presses Universitaires du Septentrion, 41-68.
- SACKS, Harvey, SCHEGLOFF, Emanuel A. & JEFFERSON, Gail (1974). A simplest systematics for the organization of turn taking for conversation, *Language*, 50: 696-735.
- SCHEGLOFF, Emanuel A. (1987). Recycled turn beginnings: a precise repair mechanism in conversation's turn-taking organization, in *Talk and social interaction*, Button, Graham & Lee, John R.E. (Eds). Clevedon : Multilingual Matters, 70-85.
- SCHEGLOFF, Emanuel A. (1996). Turn organization : one intersection of grammar and interaction, in *Interaction and grammar*, Ochs, Elinor *et al.* (Eds). Cambridge : Cambridge University Press, 54-133.

LES PROCÉDÉS D'AUTOCOPIE DE L'ENFANT APPRENTI SCRIPTEUR

Jocelyne CUSSAC POMEL, Geneviève GARCIA BOURGADE

jocelyne.pomel@wanadoo.fr, gene.bourgade@neuf.fr

CREFI-EURED, Université Toulouse Le Mirail 2,

31058 Toulouse Cedex 1

Abstract : This study looks at the spelling processes of 4-6 year old children in three word-spelling situations: list, line and separate sheets. The results show a visual proximity effect, which leads to the child copying their own work. In other words, when the child can see what they have already written, they use it as a sort of unit of measurement in order to write the next words. They copy series of letters, changing as need be the position of these letters.

1. Problématique

Quand le jeune enfant s'intéresse à l'écrit, il n'en soupçonne pas d'emblée la dimension phonographique. Il n'en traite ainsi que les propriétés à sa portée : la dimension visuelle (comment c'est écrit ?) et la dimension sémantique (qu'est-ce qui est écrit ?). Pour lui, l'écriture est un objet qui rend visible la pensée, mais qui n'est pas encore lisible.

Dans les écritures inventées, les chercheurs ont identifié un certain nombre de caractéristiques qui évoluent au fur et à mesure des bonds conceptuels que l'enfant réalise. D'abord, la quantité et la nature des lettres sont aléatoires et non systématiques, avant de se constituer en routines (Chomsky, 1975 ; Read 1986 ; Ferreiro, 1988 ; Jaffré, 1989, 1995 ; Besse, 1988 et Fijalkow & Fijalkow, 1991). Puis, l'enfant différencie objectivement ses graphies en changeant la nature et l'ordre des lettres et fait un traitement sémiographique, ajustant la quantité de lettres au référent (Ferreiro et Gomez Palaccio, 1988; Albert, 1996; Besse, 2000). L'étude quantitative de Ferreiro et Gomez Palaccio (1988) montre également que les enfants produisent une quantité minimale de lettres: au moins trois, avec une tolérance possible de un (en plus ou en moins). Autrement dit, pour l'enfant, une seule lettre ne suffit pas pour écrire un mot. Plus tard, dans son développement, l'enfant juge nécessaire non seulement un nombre minimum de lettres, mais aussi un nombre maximum (Ferreiro 2000). On observe encore que l'enfant utilise d'abord la forme de son prénom comme base graphique (Treiman, Kessler et Bourasse, 2001) avant d'étendre son répertoire à d'autres formes mémorisées.

Cependant, d'autres observations demeurent non expliquées, comme des productions dont les graphies successives ont tendance à s'allonger, le nombre de lettres augmentant d'un mot à l'autre.

2. Questionnement

Le courant piagétien (Ferreiro et Gomez Palaccio, 1988) propose une explication psychogénétique du développement de l'écriture, cependant que d'autres études explorent l'effet contexte. Par exemple, Zerbato-Poudou (2002) montre qu'en maternelle, l'enseignement de l'écriture privilégie la copie de modèle, déconnectée des apprentissages conceptuels. Nous faisons l'hypothèse que la quantité et la nature des lettres que l'enfant utilise sont en partie dépendantes du contexte et plus particulièrement du voisinage visuel. Partant du postulat que les graphies des enfants ne sont pas composées de suites de lettres totalement aléatoires, nous nous demandons quels peuvent être leurs points d'appui et les procédés qu'ils mobilisent. En ce qui concerne l'enfant qui fait un traitement visuel de

l'écrit, on peut penser qu'un de ses procédés d'écriture consiste à s'appuyer uniquement sur l'écrit antérieur qu'il a sous les yeux, conformément à une démarche reposant sur l'information visuelle.

Cette étude vise donc à observer de quelle manière les enfants marquent la différence qu'ils perçoivent entre les mots, et s'ils changent de procédé quand la tâche change. Par exemple, quand l'enfant utilise les mêmes lettres dans un ordre différent pour écrire deux mots différents (variation interfigurale), est-ce parce que ce procédé correspond à une étape génétique du développement de l'écriture ou est-ce parce que les deux mots sont écrits en liste, l'un en dessous de l'autre, en voisinage visuel?

Nous élaborons donc un plan expérimental pour étudier les relations entre les procédés visuels et la nature de la tâche, à partir des questions suivantes:

- Question 1 - Quand l'enfant peut voir les mots qu'il vient d'écrire, le mot suivant est-il écrit systématiquement avec des lettres qu'il a déjà écrites?
- Question 2 - Quand l'enfant peut voir les mots qu'il vient d'écrire, le mot suivant est-il écrit systématiquement avec plus ou moins de lettres?
- Question 3 - L'axe vertical facilite-t-il le procédé d'autocopie pour la nature et/ou la quantité des lettres reproduites?

3. Méthodologie

3.1. Echantillon

Cette étude concerne des enfants ayant délaissé le traitement figuratif de l'écrit pour en faire un traitement visuel. Nous constituons l'échantillon à l'aide d'un pré-test (écriture inventée de 4 mots) et retenons 117 enfants de 4 et 5 ans (sur 176), scolarisés en moyenne et grande section de maternelle (MS et GS), dans trois écoles (en milieu semi-urbain et urbain).

3.2. La tâche d'écriture inventée

La situation d'écriture inventée consiste à demander à l'enfant d'écrire sous la dictée quatre noms d'animaux qui ne lui ont pas été enseignés, en faisant comme s'«il savait écrire». Nous lui demandons d'identifier l'animal (verbalisation) et d'écrire son nom «*Écris lapin*». L'observation de l'enfant en train d'écrire (ses regards, son autolangage et ses paroles) nous permet de préciser s'il utilise l'autocopie. Un entretien minimum complète l'observation quand le procédé utilisé par l'enfant n'est pas clairement identifié, notamment en ce qui concerne la valeur sonore des lettres. Pour trois enfants, une étude longitudinale de 2 mois est menée, avec des passations plus nombreuses et des entretiens cliniques enregistrés (série de questions sur la nature et la quantité des lettres). Ces données quantitatives sont destinées à expliquer les procédés observés.

3.3. Conditions expérimentales

Trois tâches sont définies: l'écriture sur des feuilles différentes, qui exclue l'autocopie, l'écriture en liste, supposée favoriser l'autocopie et l'écriture en ligne, supposée médiane entre les deux tâches précédentes, n'empêchant pas l'autocopie et ne la facilitant pas non plus. A ces trois tâches, correspondent trois conditions expérimentales:

1. L'enfant écrit une série de 4 mots, en liste sur une même feuille «les uns en dessous des autres».
2. Il écrit une série de 4 mots, en ligne sur une même feuille «les uns à côté des autres».
3. Il écrit une série de 4 mots, avec une feuille différente pour chaque mot.

Les 117 enfants sont répartis en 3 groupes de 26 GS et 13 MS. Chaque groupe effectue, à quelques jours d'intervalle, deux des trois tâches suivantes: écrire 4 mots bi-syllabiques en liste, en ligne et sur des feuilles différentes (tableau 1). L'ordre de passation des tâches et celui de l'écriture des mots sont aléatoires. Les passations ont lieu en petits groupes de 2 à 4

Tâche	Mots en ligne	Mots isolés	enfants pour favoriser le
	condition expérimentale 1	condition expérimentale 3	langage égocentrique
	liste / isolés	isolés / ligne	(Piaget, 1936 ; Vygotski,
Mots en liste	26 GS et 13 MS	26 GS et 13 MS	1934) et minorer l'effet
	condition expérimentale 2		adulte (Milgram, 1974).
	liste / ligne		Nous recherchons la
	26 GS et 13 MS		signification des
Tableau 1 - Plan expérimental			différences de moyenne

à l'aide d'une ANOVA (Statview) à partir des variables suivantes : pour la nature des lettres, le nombre de lettres pour chaque suite reprise, et pour la quantité de lettres, d'une part, le nombre de lettres de la graphie et sa longueur physique, et, d'autre part, les régularités (quantités égales, croissantes et décroissantes).

4. Résultats

A l'issue du plan expérimental, nous disposons d'un échantillon de 99 enfants de 4 à 6 ans, de MS et GS. L'analyse des productions met en évidence un certain nombre de procédés d'écriture qui concernent aussi bien la nature et la position des lettres que leur quantité. Certains de ces procédés visuels relèvent d'un comportement d'autocopie et sont favorisés par la nature de la tâche. Pour être autonomes sur la tâche demandée, les enfants se donnent leurs propres règles d'écriture. Ils exploitent le voisinage visuel, plus spécialement quand ils écrivent en liste, pour copier leur propre production.

4.1. Question 1

Quand l'enfant peut voir les mots qu'il vient d'écrire, le mot suivant est-il écrit systématiquement avec des lettres qu'il a déjà écrites?

Pour différencier l'autocopie de la simple reprise de lettres (inhérente à toute production), nous observons les reprises de suite de lettres. Comme il est plus facile de mémoriser une suite de 2 lettres qu'une suite de 3, etc., nous nous attendons à observer davantage de reprises de 3 lettres que de reprises de 2 lettres, dans les tâches avec voisinage visuel (liste et ligne), davantage de reprises de 4 lettres que de reprises de 3 lettres dans ces mêmes tâches.

Ce qui revient à dire que nous nous attendons à ce que ces tâches avec voisinage visuel favorisent l'autocopie.

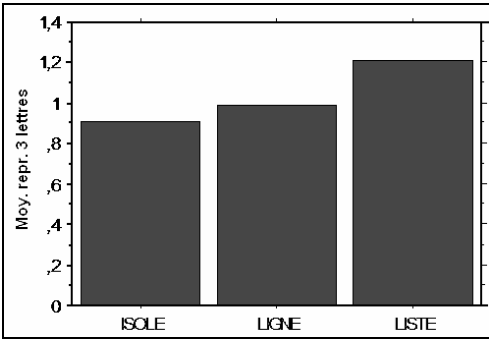


Figure 1 - Nombre de reprises d'au moins 3 lettres

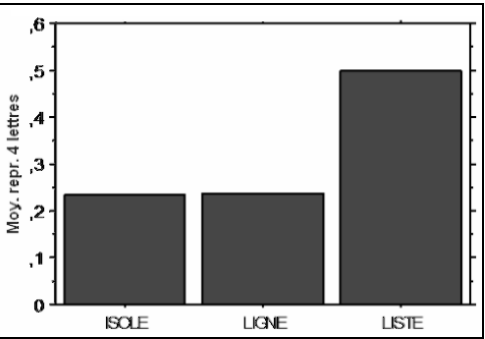


Figure 2 - Nombre de reprise d'au moins 4 lettres

Quel que soit le seuil choisi pour le nombre de lettres reprises, une tendance générale se dessine en faveur de l'écriture en «liste», suivie de l'écriture en «ligne»; l'écriture en «isolé» présentant les scores les plus bas (au -2 lettres, moy.=1,61 liste, =1,46 ligne, =1,44 isolé; au -3 lettres, moy.=1,21 liste, =0,98 ligne, =0,91 isolé; au -4 lettres, moy.=0,50 liste, =0,24 ligne, =0,23 isolé; au -5 lettres, moy.=0,16 liste, =0,079 ligne, =0,078 isolé) (Figures 1 et 2).

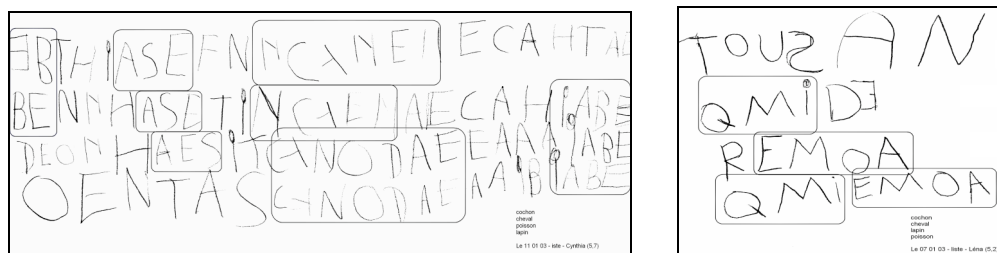
Dans les tâches avec voisinage visuel (liste et ligne), l'enfant a tendance à reprendre plus souvent des suites de lettres déjà écrites que dans la tâche sans voisinage visuel (isolé). L'enfant s'appuierait donc sur le voisinage visuel pour copier des suites de lettres déjà utilisées.

On observe de plus que les différences de moyenne entre «ligne» et «liste» ont tendance à augmenter au fur et à mesure que l'on considère un plus grand nombre de lettres, jusqu'au seuil de «reprise d'au moins 4 lettres» (au -2 lettres, $D=0,15$; au -3 lettres, $D=0,23$; au -4 lettres, $D=0,26$; au -5 lettres, $D=0,08$) et que les différences de moyenne entre «ligne» et «isolé» ont tendance à diminuer au fur et à mesure que l'on considère un plus grand nombre de lettres, à partir du seuil d'au -3 lettres (au -2 lettres, $D=0,014$; au -3 lettres, $D=0,07$; au -4 lettres, $D=0,004$; au -5 lettres, $D=0,002$). Ainsi, plus le nombre de lettres de la suite augmente, plus le comportement de l'enfant change: plus il se comporte de manière identique en «ligne» et en «isolé».

On observe enfin une différence significative ($F(2, N=4,060)$; $p<0,05$) entre l'écriture en liste et les deux autres tâches, à ce même seuil de «reprise d'au moins 4 lettres», ce qui montre que l'écriture en liste est la seule tâche des trois qui favorise la reprise d'au moins 4 lettres.

Les reprises de suites de lettres en liste peuvent donc être interprétées comme un procédé d'autocopie que l'enfant mobilise quand le contexte le lui permet.

Les productions montrent des procédés différents selon le niveau de traitement de l'écrit. Par exemple, Cynthia (5,7), qui écrit des lignes entières de lettres pour un mot, copie dans l'axe vertical des suites de 2 à 6 lettres quand elle écrit en liste (Figure 3); alors que Léna (5,2) nous présente un procédé plus complexe (Figure 4). Ses graphies ont une quantité de lettres qui indiquent qu'elle a une représentation plus accomplie de ce que peut être un mot. Elle profite du voisinage visuel pour fabriquer un mot uniquement avec des suites de lettres prélevées dans ses graphies précédentes. On peut voir qu'elle reprend ainsi la suite de lettres QMI de la 2^{ème} graphie et la suite EMOA de la 3^{ème} pour former la 4^{ème} graphie.



Figures 3 et 4 - Autocopie de suites de lettres

4.2. Question 2

Quand l'enfant peut voir les mots qu'il vient d'écrire, le mot suivant est-il écrit systématiquement avec plus ou moins de lettres?

Les procédés d'autocopie se rapportant à la quantité de lettres que nous avons observés sont très variés. Certains enfants font référence au nombre de lettres et d'autres à la longueur physique de la graphie. De même, certains enfants expliquent vouloir obtenir la même quantité de lettres, et d'autres, au contraire «faire plus long» ou «pas pareil». Cette diversité

de procédés nous amène donc à observer des croisements de variables: d'une part, les écarts type du nombre de lettres et de la longueur de la graphie (+ ou - 0,5 mm); d'autre part, les régularités suivantes d'un mot à l'autre: quantités égales (ou quasi-égales), croissantes et décroissantes.

4.2.1. Quantités égales

Quand les graphies présentent une même quantité de lettres (ou presque), cette quantité peut témoigner d'un traitement de l'oral (les mots proposés à l'écriture étant des mots bisyllabiques), d'une référence à une quantité de lettres mémorisée ou d'un comportement d'autocopie. Nous nous attendons donc à trouver un écart type plus important dans la tâche sans voisinage visuel (isolé) que dans les tâches avec voisinage visuel (liste et ligne).

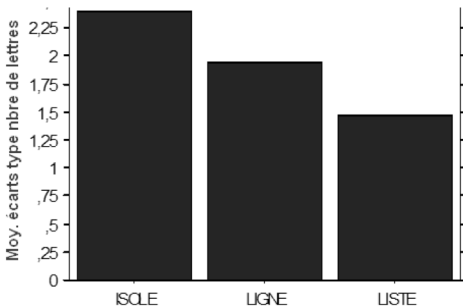


Figure 5 - Ecart-type du nombre de lettres

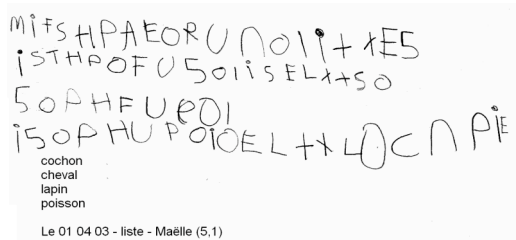
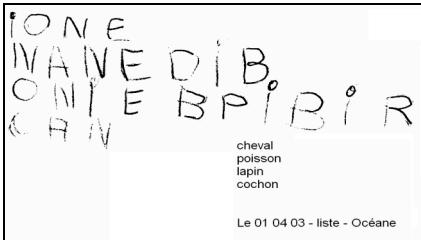
pour les « vieux ». Notre interprétation est que l'écriture en liste offre le voisinage visuel des graphies, et de ce fait, exerce plus d'influence sur l'écriture des plus jeunes enfants, ce qui se traduit par des graphies qui ont tendance à se ressembler en termes de nombre de lettres.

4.2.2. Quantités croissantes et décroissantes

Nous relevons le nombre de relations croissantes et décroissantes d'un mot à l'autre et nous nous attendons à observer une différence selon que l'enfant est en voisinage visuel (liste et ligne) ou pas (isolé).

Bien que certaines productions laissent penser à des procédés d'autocopie de ce type, nous n'observons pas de différence statistique significative entre les trois tâches, quant à une tendance des enfants à allonger ou diminuer leurs graphies.

Cependant, ces procédés d'autocopie se rencontrent et les entretiens métagraphiques révèlent que les graphies sont bien écrites en différenciation de quantité les unes par rapport aux autres.



Figures 6 et 7 - Unité de mesure (1) et (2)

Par exemple, Océane se donne une règle d'alternance de longueur. Elle a écrit quatre mots en liste: un court, un long, un plus long et un court (Figure 6). Elle explique ainsi les différences de longueur:

«Adulte. -Tu as fait exprès de l'écrire plus long ? / Océane. - Oui, parce que sinon on va pas séparer les deux [mots] et il faut pas que les deux petits [mots] ils se touchent» (Océane, 5,10)

Maëlle, elle, use de la métaphore de «Boucle d'Or et les trois ours» pour expliquer qu'elle différencie ses graphies par leur longueur (Figure 7):

«Adulte -Et celui-là, il est...? / Maëlle- Petit. Parce que moi je fais les toujours le gros, le moyen et le petit. / A. -Tu fais toujours comme ça? / M. - Parce que j'ai envie de continuer l'histoire presque. Parce que ça c'est le papa ours, ça c'est le maman ours et le petit. / A. - Oui, mais moi je vous ai demandé d'écrire cochon, cheval... / M. - Mais je l'ai écrit! En fait j'ai pensé dans ma tête à faire ça, mais j'ai écrit cochon, cheval, lapin et aussi dans ma tête j'ai pensé à une idée que ça ce soit le papa, ça la maman et ça le petit.» (Maëlle, 5,1)

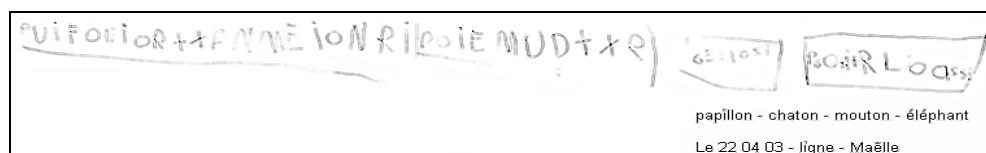


Figure 8 - Unité de mesure (3)

Maëlle est séduite par ce procédé, si bien qu'elle l'utilise encore pour écrire d'autres mots, une vingtaine de jours plus tard, cette fois, en ligne (Figure 8). L'entretien nous montre que Maëlle est contrariée par la nouvelle consigne d'écrire les mots en ligne, elle sent bien qu'il va lui être difficile de mettre en œuvre son procédé favori. Quand elle a terminé l'écriture du premier mot, elle annonce:

«M. - J'ai fait très long./ Adulte – Pourquoi? / M. – Parce que j'ai envie de faire gros, gros et tout petit. / A. – Il faut l'écrire à côté de l'autre. / M. – Moi je préfère l'écrire en bas. / A. – Non, moi je veux que tu fasses une ligne avec ces mots. / M. - Alors tant pis ça va être pareil, je vais écrire le gros, ça va être pareil. Je vais écrire le moyen, ça veut dire pas très grand, ça va être pareil.» (Maëlle, 5,1).

Pour conclure, quand l'enfant peut voir les mots qu'il vient d'écrire, le mot suivant n'est pas écrit systématiquement avec plus ou moins de lettres. Par contre, quand l'enfant écrit en liste, le mot suivant a tendance à être écrit avec le même nombre de lettres.

4.3. Question 3

L'axe vertical facilite-t-il le procédé d'autocopie pour la nature et/ou la quantité des lettres reproduites?

En regard des résultats précédents, l'axe vertical (écriture en liste) favorise les procédés d'autocopie en ce qui concerne la nature des lettres utilisées. Quand les mots sont écrits les uns en dessous des autres (en liste), le mot suivant est bien écrit systématiquement avec des lettres déjà écrites, mais pas quand ils sont écrits les uns à côté des autres (en ligne).

5. Conclusion

Ainsi, le voisinage visuel semble bien avoir un effet sur l'écriture inventée des enfants en ce qui concerne la nature et la quantité des lettres utilisées. Parce qu'il est dans une tâche donnée, l'enfant organise la réalisation de la tâche en fonction du contexte. L'écriture en liste favorise les copies de suites de 3 ou 4 lettres et la copie d'un même nombre de lettres, surtout pour les enfants les plus jeunes. L'écriture en ligne apparaît comme une tâche intermédiaire entre «isolé» et «liste». Elle favorise certains comportements d'autocopie, mais dans une moindre mesure que l'écriture en liste.

Très souvent, les procédés d'autocopie se combinent dans les productions. Par exemple, Steve utilise les mêmes lettres pour écrire les deux premiers mots, puis tourne avec d'autres lettres pour les deux derniers mots et, bien que les quatre mots aient le même nombre de lettres, la différence de taille des lettres crée l'illusion de mots plus longs (Figure 9).

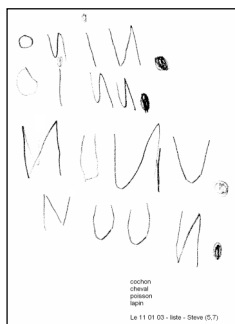


Figure 9 -
Combinaison de pro-
cédés d'autocopie

Ces résultats peuvent s'expliquer si l'on considère que le champ visuel d'un enfant de 4/6 ans est plus réduit que celui de l'adulte, et qu'il prend plus facilement en compte ce qui se trouve en face de lui, que ce qu'il y a sur le côté. Hors, c'est précisément quand l'enfant écrit en liste, qu'il conserve dans son champ visuel ce qu'il a déjà écrit, contrairement aux deux autres tâches. Quand il écrit chaque mot isolément, sur des feuilles différentes, l'enfant ne voit plus ses graphies précédentes et quand il écrit en ligne, les graphies précédentes sortent de son champ de vision à mesure qu'il écrit. De plus, de nombreux enfants (22 pour 63) tracent des signes si grands et des graphies si longues qu'ils utilisent un support de deux à quatre feuilles quand ils écrivent en ligne.

On observe, enfin, que ces procédés d'autocopie accompagnent les changements de conceptualisation des enfants.

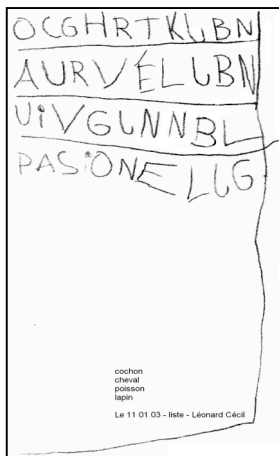


Figure 10 - Traitement oral / autocopie (1)



Figure 11 - Traitement oral / autocopie (2)

Par exemple, Léonard Cécil (5,6) trace un cadre dans lequel il inscrit exactement tous les mots, ce qui leur donne la même longueur physique, bien que le nombre de lettres varie (Figure 10). On observe qu'il commence tout juste à coder l'oral sur la première lettre de certains mots. Trois mois plus tard, il a basculé dans un traitement phonémique tout en continuant à tracer un cadre de référence. On peut voir sur sa production les deux types de traitement (oral et visuel) se confronter (Figure 11). Léonard Cécil s'autorise parfois à sortir du cadre, comme pour le E final

de «cheval»; et parfois il ne s'y autorise pas, comme pour le INE de «lapin» qu'il préfère écrire en faisant un retour à la ligne.

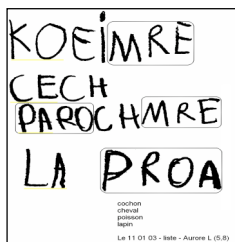


Figure 12 -
Traitement oral /
autocopie (3)

Aurore (5,8) nous montre un autre procédé d'autocopie destiné, sans doute, à coder l'oral (Figure 12). Elle fait un traitement phonémique au début des mots (cochon, cheval, poisson, lapin) et termine ses graphies avec des reprises de suites de lettres : il semble qu'elle reprenne MRE pour coder le même son [ø] et qu'elle permute PARO en PROA pour signifier l'analogie qu'elle repère entre [pwa] et [pẽ].

Ces résultats plaident en premier lieu, en faveur d'un processus cognitif d'entrée dans l'écrit plus contextuel et individuel qu'étapiste.

De plus, ils permettent de conclure à la nécessité d'une grande prudence méthodologique dans l'analyse des écritures inventées.

Dans une tâche de ce type, l'écriture de l'enfant s'y déploie dans un mouvement dynamique entre ce qu'il sait de l'écrit et ce qu'il est en train de produire en

contexte et certains enfants se révèlent, plus que d'autres, dépendants de la nature de la tâche demandée.

Enfin, que les productions enfantines varient autant quand le contexte varie, relativise l'explication psychogénétique du développement de l'écriture, notamment en ce qui concerne les notions de «nombre d'or», de «quantité minimale» et «maximale». Certains enfants écrivent bien avec un nombre fixe de lettres, mais ce nombre varie d'une tâche à l'autre et d'une passation à l'autre. Certains enfants utilisent, en alternance, d'une passation à l'autre, toute une ligne de lettres pour un mot et un nombre limité de lettres. L'explication d'un effet plafond n'est pas satisfaisante dans ce cas. L'enfant peut ne pas montrer d'effet plafond, mais un effet que nous qualifions plutôt de psychoaffectif: l'enfant trace pour le plaisir et oublie la consigne.

Enfin, la production d'Aurore montre une combinaison de deux types de procédés: l'un est visuel (l'autocopie), l'autre est oral (traitement phonémique de la première syllabe) (Figure 12). Nous avons mené une expérimentation didactique en cours d'analyse qui étudie comment et quand les enfants abandonnent leurs procédés d'autocopie au profit d'un traitement oral du mot.

Références bibliographiques

- ALBERT, Jacqy (1996). La construction de l'écrit par les enfants scolarisés en 1^{ère} année du cycle des apprentissages fondamentaux, *Psychologie et éducation*, 24: 57-73.
- BESSE, Jean-Marie (1988). *L'enfant et la construction de la langue écrite: une approche génétique*. Lyon : Voies livres 9.
- BESSE, Jean-Marie (2000). *Regarde comme j'écris! Ecrits d'élève, regards d'enseignants*. Paris: Magnard.
- CHOMSKY, Carol (1975). Invented spelling in the open classroom, *Child Language Today*, numéro spécial: 400-518.
- FERREIRO, Emilia (2000). *L'écriture avant la lettre*. Paris : Hachette éducation.
- FERREIRO, Emilia, GOMEZ PALACIO, Margarita (1988). *Lire-écrire à l'école comment s'y apprennent-ils ? : Analyse des perturbations dans les processus d'apprentissage de la lecture et de l'écriture*. Lyon : CRDP.
- FIJALKOW, Jacques & FIJALKOW, Eliane (1991). L'écriture inventée au cycle des apprentissages, *Les Dossiers de l'éducation*, 18: 125-167.
- JAFFRE, Jean-Pierre (1989). La découverte de l'écriture : aspects didactiques et psycholinguistiques, *Premier contact avec l'écriture et la lecture*. Grèce: Hélidoni.
- JAFFRE, Jean-Pierre (1995). Le traitement élémentaire de l'orthographe: les procédures graphiques, *Langue française*, 95 : 27-48.
- MILGRAM, Stanley (1974). *Soumission à l'autorité*. Paris: Calmann-Lévy.
- PIAGET, Jean (1936). *La naissance de l'intelligence chez l'enfant*. Neuchâtel-Paris: Delachaux et Niestlé, 9^{ème} édition 1977.
- READ, Charles (1986). *Children's creative spelling*. London: Routledge and Kegan Paul.
- TREIMAN, Rebecca, KESSLER, Brett & BOURASSA, Derrick (2001). Children's own names influence their spelling, *Applied Psycholinguistics*, 22: 555-570.
- VYGOTSKI, Lev Semionovitch (1985, 1^{ère} édition 1934). *Pensée et langage*. Trad. fr., Paris : Messidor/Éditions Sociales.
- ZERBATO-POUDOU, Marie-Thérèse (2002). A quoi servent les exercices graphiques? L'écriture et son apprentissage à l'école élémentaire, *Repères* n°26/27 : 61-71.

COMMENT DIRE OUI ? LES MARQUES D'ACCORD DANS LE CORPUS LANCOM

ANALYSE DIDACTIQUE ET LINGUISTIQUE

Juliette DELAHAIE

jdelahaie@hotmail.com

Laboratoire MoDyCo

92001 Paris-X-Nanterre

Abstract : LanCom is a native/non-native corpus of spoken French interactions. Its analysis shows that non-native learners of French as a foreign language and French native speakers express agreement in very different ways. This paper is a linguistic study of some terms of agreement, like *bien sûr*, *d'accord*, or *voilà*, which native speakers frequently use in LanCom, whereas non-native speakers do not. Related teaching applications are also discussed.

Introduction

Comment dire oui ? La réponse semble aller de soi, le français possédant un certain nombre de marques d'accord, comme *oui*, *d'accord*, *voilà*, *tout à fait*, etc. qui sont dévolues à cette fonction. Or, quand on analyse de près ce que font des apprenants non-natifs de français langue étrangère quand ils conversent en français, on s'aperçoit qu'au contraire leur utilisation pose un certain nombre de problèmes : on entend un *bien sûr* là où l'on attendrait plutôt un *d'accord*, un *oui* là où l'on aurait attendu un renchérissement par *voilà*. Bref, le maniement des marques d'accord n'est pas si simple, et il est relativement important dans la mesure où il participe à la politesse conversationnelle qui, sous nos latitudes, privilégie le consensus et la coopération interactionnelle¹. Les professionnels du français langue étrangère encouragent d'ailleurs à enseigner les stratégies interactives de « coopération interpersonnelle » qui permettent de « commencer, soutenir et terminer une conversation avec naturel et avec des tours de parole efficaces » (Cadre européen commun de référence pour les langues, 2001 : 69-70). Les marques d'accord pourraient facilement rentrer dans ce cadre.

Difficiles à manier, les marques d'accord sont aussi difficiles à cerner sur le plan définitionnel. Grammaticalement parlant, il ne s'agit pas d'une catégorie homogène : un grand nombre de marques relève de la classe des adverbes en tant que « marqueurs des actes illocutoires de l'acquiescement » (Riegel, Rioul & Pellat, 2002 : 376), d'autres comme *hm hm* sont plutôt classées dans les interjections. L'expression « marques d'accord » permet de dépasser cette hétérogénéité grammaticale et de désigner les mots qui vont de l'expression de la régulation à l'expression d'un accord dans un tour de parole en bonne et due forme, comme réponse ou enchaînement à toute une série d'actes initiatifs tels que la question, l'assertion, la demande de confirmation ou l'offre. En analyse interactionnelle, on peut également leur trouver des affinités avec ces « petits mots » qui participent au maintien du flux discursif par « une activité minimum de gestion assurant la continuité du fil tissé par l'interlocuteur » (Bruxelles & Traverso, 2001 : 47). Mais, parmi ces derniers, les chercheurs se sont plutôt intéressés à d'autres marques qui n'expriment pas directement l'accord comme *ben* (Bruxelles & Traverso, 2001), *hein* (Delomier, 1999), *quoi* (Chanet, 2001).

¹ Voir à ce propos Kerbrat-Orecchioni, 1994.

Les marques d'accord constituent donc un ensemble ouvert et polymorphe que le champ de la linguistique (traditionnelle, discursive, interactionnelle) n'a pas encore entièrement investi, peut-être parce qu'elles ne posent pas *a priori* de problème particulier. Ce sont les difficultés et hésitations du non-natif qui vont du même coup attirer l'attention du natif sur des faits de langue passés inaperçus. L'objectif sera ici à la fois linguistique et didactique : il s'agira de proposer quelques explications linguistiques et interactionnelles de quelques marques d'accord qui posent problème aux apprenants de français langue étrangère, de façon à en faciliter l'apprentissage en classe.

Pour cela, nous partirons de données concrètes tirées du corpus LanCom, corpus différentiel natif/non-natif élaboré à partir des années 1990 en Belgique néerlandophone par l'Université de Louvain-Leuven. Ce corpus accessible sur Internet est composé de deux volets : une partie a été enregistrée dans les classes de l'enseignement secondaire en Flandre, où les élèves ont joué sous forme de jeux de rôle des situations de la vie quotidienne, l'autre partie est composée des mêmes situations jouées ou filmées de manière authentique en France. Les données que nous utiliserons ici constituent une sorte de sous-corpus de 19 645 mots, formé de cinq transactions dans une agence de voyage à Lille (filmées en février 1993 et en juin 2006), avec trois employées différentes, et de treize jeux de rôle du même type de transaction joués par trois classes du secondaire en Belgique néerlandophone (avril 2006). Le type d'interaction qu'induit une agence de voyage est particulièrement intéressant pour l'étude des marques d'accord : dans ces interactions relativement codifiées, les interlocuteurs qui ont des rôles très spécialisés, celui de vendeur et de client, doivent mettre en place tout un système de coopération interpersonnelle afin de parvenir au terme de l'échange qui est l'obtention du bien ou sa promesse.

Or la coopération n'est pas exprimée de la même manière chez francophones et néerlandophones, c'est ce que nous montrera l'étude comparative de l'ensemble des marques mobilisées par les deux groupes. Si le néerlandophone exprime bien son accord, il le fait tout autrement, pourquoi ? On se demandera dans un second temps s'il ne serait pas influencé dans ses choix par l'input éducationnel, c'est-à-dire le discours du professeur et les manuels de français langue étrangère. Enfin, on tentera de dégager quelques traits marquants de quelques marques d'accord qui posent problème aux non-francophones et que ne détaillent pas les manuels.

1. Etude comparative : deux manières différentes de marquer l'accord

Comment des apprenants néerlandophones de français langue étrangère (FLE) expriment-ils l'accord ? L'hypothèse était que des apprenants de niveau avancé se rapprocheraient de l'habileté du natif : en 4^e, 5^e et 6^e de l'enseignement secondaire en Flandre, les apprenants enregistrés ont entre quinze et dix-sept ans, et étudient le français depuis l'âge de dix ans, à raison de quatre à cinq heures par semaine. On peut donc s'attendre à ce qu'ils maîtrisent plus ou moins bien la conversation en français et ses marques d'accord, notamment en classe de 6^e. Or, de la 4^e à la 6^e, il y a bien une évolution au niveau de la maîtrise grammaticale et lexicale ainsi que de l'aisance à communiquer, mais pas de changement en ce qui concerne les marques d'accord. Toutes les interactions néerlandophones ont donc été regroupées et comparées à ce que font les natifs dans la même situation.

Au niveau méthodologique, on a rangé dans la catégorie marques d'accord les mots qui appartiennent en général à la catégorie grammaticale d'adverbe ou d'interjection (à part *voilà* qui est un « présentatif »), et on a écarté de notre analyse les expressions d'accord, à part *c'est ça*, *c'est bien* et *c'est bon*, parce que leur sens ne semble pas véhiculer plus qu'un simple accord. On a comptabilisé les occurrences de chaque marque, sans tenir compte des associations possibles et souvent réalisées entre elles, ce qui nécessiterait une étude à part

entière. Le graphique suivant permet de visualiser et de chiffrer les différences qui existent entre la pratique de l'apprenant néerlandophone et celle du francophone :

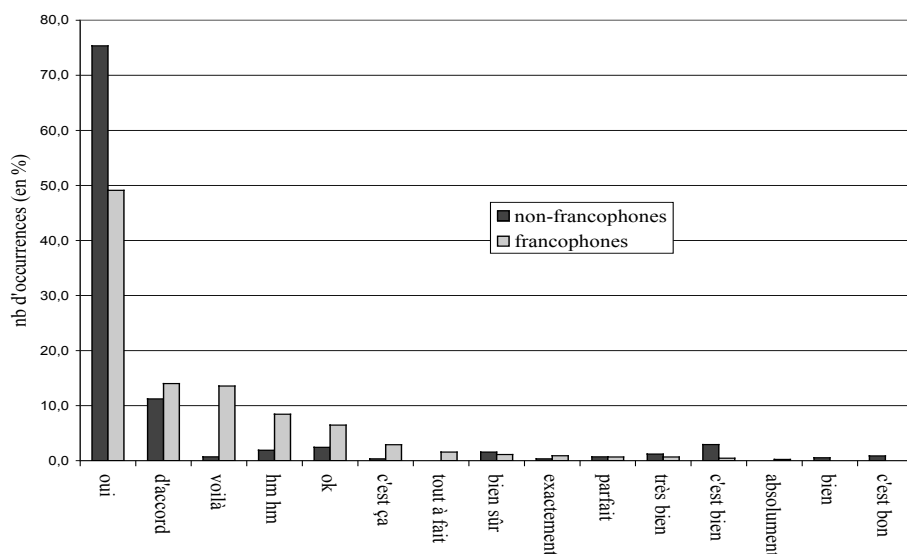


Figure 1 - Francophones et non-francophones : deux manières différentes de marquer l'accord

Sur 10 000 mots dans le volet néerlandophone, on a comptabilisé 580 marques d'accord, contre 450 dans le volet francophone, pour 9 645 mots. Autrement dit, le néerlandophone acquiesce légèrement plus que le francophone (5,8 marques d'accord pour 100 mots contre 4,6 chez le francophone), mais il le fait d'une manière bien différente. Tout d'abord, cette plus forte proportion de marques d'accord chez le néerlandophone est due à une forte proportion de *oui* (75,3% de *oui*, contre 49,1% chez le francophone). Ensuite, par rapport aux francophones, les apprenants néerlandophones sous-emploient certaines marques comme *voilà* (13,6 % chez les francophones, 0,7% chez les apprenants néerlandophones), *ok* (6,4% contre 2,4%), *c'est ça* (2,9% contre 0,3%). D'autres marques sont totalement absentes de leur répertoire, comme *tout à fait* (1,6% chez le francophone) et *absolument* (0,5%). Enfin, les apprenants néerlandophones emploient des marques peu ou pas employées par les francophones comme *c'est bien* (2,9% chez le néerlandophone contre 0,4% chez le francophone) et *c'est bon* (0,9% chez le néerlandophone) dans des échanges du type² :

E1- *ah je ne sais pas pas pas certainement donc je : je vais faire un coup de téléphone*

E2- *ah c'est bon ça (jeu de rôle ; 5e n°2)*

E1- *bonjour ici c'est (prénom nom) de : de l'a l'agence de voyage de Diksmuide euh est-ce qu'il y a encore une voiture de type A libre dans la période de trois août au euh du trois août au huit euh au dix août [...]*

E2- *euh oui c'est bien nous avons encore une voiture (jeu de rôle ; 5e n°3)*

Ces deux expressions sonnent étrangement à l'oreille du francophone qui aurait plutôt employé la locution *d'accord*. Non attestées en français de Belgique (selon une informatrice belge), ces marques sont la traduction littérale de l'expression *het is goed* qui est l'équivalent de *d'accord* en flamand (selon un informateur flamand).

² Dans les transcriptions, E1 et E2 désignent des élèves.

En conclusion, cette étude comparée de la fréquence et du type des marques d'accord dans les scènes d'agence de voyage montre que le néerlandophone acquiesce, et même plus que le francophone, ce qui va à l'encontre de certaines idées reçues selon lesquelles les locuteurs des pays du nord seraient moins coopératifs dans les interactions et avant tout orientés vers le but, tandis que leurs pairs latins seraient très attentifs à la relation avec leurs interlocuteurs (voir par exemple Hall et Hall, 1990 : 49). Il semblerait qu'au contraire l'apprenant essaye de marquer à toutes les étapes de l'interaction sa volonté de coopération et sa compréhension. Cependant, il ne le fait pas de la même manière que le francophone, pourquoi ? On se demandera si ce contraste ne serait pas lié à la manière dont sont enseignées les marques d'accord, à la fois chez l'enseignant et dans les manuels de FLE.

2. L'enseignement des marques d'accord : la « malédiction paradigmatique »³

2.1. L'enseignant : combler les silences et enrichir le vocabulaire

Les évaluations réalisées par les enseignants ont été enregistrées à la suite de chaque jeu de rôle. Parmi eux, une seule enseignante engage explicitement ses élèves à marquer l'accord par des expressions diverses⁴ :

P- oui c'est très bien n'oubliez pas hein de épeler = n'oubliez pas non plus les interjections au téléphone = il y a des groupes qui font ça très très bien quelles sont les interjections au téléphone ? pas tout tout le temps hm hm hm hm qu'est-ce qu'on peut dire d'autre ? = aucune idée ? oui

E- d'accord

P- d'accord = oui ? (prénom) ?

E- bien

P- bien = autre chose ?

E- je vois

P- je vois bien entendu je comprends c'est parfait et cætera hein donc il faut compléter il faut meubler un petit peu les silences [...] (4e, évaluation au jeu de rôle)

L'enseignante attire l'attention sur les interjections qui permettent de faire vrai au téléphone et qui sont en réalité des marques ou expressions d'accord. Elle a bien saisi que le francophone répugne à laisser des silences dans la conversation et qu'il manifeste constamment son engagement par des marques d'accord. Cependant, les marques d'accord qu'elle propose ne sont pas satisfaisantes : dans les conversations enregistrées, les marques comme *je vois*, *bien entendu*, *je comprends*, *c'est parfait*, ne sont pas employées, même au téléphone. C'est que l'objectif de l'enseignante n'est pas seulement d'enseigner le français parlé : dans un cadre didactique, ces interjections participent à l'enrichissement du vocabulaire des apprenants, et à la volonté d'enseigner ainsi tout un paradigme de synonymes.

2.2. Les manuels de français langue étrangère : l'effet de liste

Les manuels de FLE, aussi bien ceux utilisés dans les classes filmées que ceux du marché français, cèdent à la même « illusion paradigmatique » : les marques d'accord sont souvent présentées sous forme de listes dans lesquelles se mêlent expressions rares et mots usuels. On donnera un exemple de liste tiré d'un manuel récent du marché français,

³ Expression due à Beacco, 1989 : 144.

⁴ Dans l'échange retranscrit, P est l'enseignante et E un élève.

Forum 2 (2001 : 77) : pour « exprimer son accord/approuver », le manuel propose la liste suivante : « c'est évident, vous avez raison, évidemment, précisément, en effet, absolument, tout à fait, bien entendu ». Dans cette liste se trouvent des expressions qui apparaissent dans notre corpus, comme *tout à fait* et *absolument*, d'autres qui sont d'un emploi très rare en français parlé en tant que marques d'accord, comme *précisément* ou *c'est évident*⁵.

En conclusion, enseignants et manuels tentent d'enseigner quelques rudiments pour l'emploi des marques d'accord, mais il y a chez les uns et les autres un problème en ce qui concerne la sélection des marques (ou plutôt son absence), et un manque de réflexion sur leurs fonctions potentiellement différentes. Une analyse attentive révèle en effet que, loin d'être interchangeables, les marques d'accord possèdent des valeurs propres. Nous tenterons ainsi de dégager la valeur de marques d'accord courantes chez le natif, qui posent problème aux apprenants et que ne détaillent pas les manuels.

3. Analyse linguistique et interactionnelle de quelques marques d'accord

Nous avons choisi des marques très courantes chez les francophones⁶, *voilà*, *d'accord* et *ok*, d'autres qui le sont moins mais qui n'apparaissent pas chez les néerlandophones, comme *tout à fait*, et enfin *bien sûr*, parce que cette marque pose des problèmes particuliers aux apprenants. L'objectif ne sera pas ici de détailler toutes les fonctions en discours de ces marques d'accord, mais de mettre en valeur leur différence du point de vue pragmatique et interactionnel : leur emploi est fortement lié au rapport de place entre les interlocuteurs, à leurs statuts (ici en l'occurrence institutionnels) et leurs rôles (vendeur vs client).

3.1. *voilà et d'accord*

C'est un couple intéressant parce qu'il peut avoir des emplois radicalement différents qui permettent de cerner en partie leur valeur respective. De plus, les néerlandophones ne connaissent pas *voilà* en tant que marque d'accord, mais l'utilisent en tant que « présentatif » d'objet ou de personne, conformément à sa définition dans les grammaires usuelles et les manuels de FLE.

Et pourtant, *voilà* est une marque d'accord très employée par les francophones. C'est un instrument de réponse privilégié aux demandes de confirmation, avec 24% de ses occurrences consacrées à cet emploi. Dans la demande de confirmation, énoncé intermédiaire entre la question et l'assertion, l'interlocuteur A affirme quelque chose concernant un fait b, et l'interlocuteur B interprétera alors cette affirmation comme une requête de confirmation (voir Kerbrat-Orecchioni, 1991 : 93-94). On illustrera cette définition par l'exemple suivant tiré du corpus francophone⁷:

C- *oui le je oui parce que lui il prend le le vol de un vol quelques heures avant je crois*

E- *voilà treize heures quarante cinq*

?? *d'accord treize heures quarante-cinq (agence de voyage, « un gentil client »)*

La demande de confirmation prend la forme de l'assertion, mais l'incertitude est marquée par l'expression modalisatrice *je crois*. Dans un tel contexte, *d'accord* est improbable, l'employée ne pouvant pas avouer qu'elle ne connaît pas son dossier et son client. Autrement

⁵ Vérification a été faite à partir de la base de données Elicop.

⁶ Voir les pourcentages du graphique.

⁷ Dans l'échange retranscrit, C : client et E : employée.

dit, le statut et le rôle de l'employé autorisent et même nécessitent la réponse en *voilà*. En revanche, c'est *voilà* qui est incorrect lorsqu'il s'agit de ratifier une réponse dans le troisième membre d'un échange que l'on appelle évaluatif, « par lequel L1 clôt cet échange qu'il a lui-même ouvert, en signalant à L2 qu'il a bien enregistré son intervention réactive, et qu'il la juge satisfaisante » (Kerbrat-Orecchioni, 1998 : 236). La marque d'accord serait donc à ranger dans la catégorie des « accusés de réception » :

E- au niveau formalités vous avez ce qu'il faut vous avez une carte d'identité ou un passeport ?

C- oui oui oui

E- d'accord (agence de voyage, « vacances à Lesbos »)

Dans l'échange suivant, il est impossible de remplacer *d'accord* par *voilà*, en revanche dans le cadre de l'enchaînement sur une (vraie) assertion, *voilà* et *d'accord* peuvent également apparaître, mais avec des valeurs différentes :

C- monsieur Freignac sera à l'accueil à l'aéroport

E- voilà il sera là

E- d'accord il sera là (agence de voyage, « un gentil client »)

Dans cet échange, l'employée peut répondre à l'assertion du client par *voilà*, et c'est d'ailleurs ce qu'elle choisit : elle est ici habilitée à le faire, car elle est censée en savoir autant si ce n'est plus que son client sur son dossier de voyage, elle peut donc confirmer l'assertion et montrer qu'elle connaît bien son dossier et son métier. L'employée peut aussi préférer user d'une marque de ratification plus neutre et laisser la supériorité de savoir au locuteur C, considérant que l'assertion a été un réel apport d'information enregistré par *d'accord*.

Voilà et *d'accord* sont donc pourvus de valeurs différentes, l'un marque la confirmation tandis que l'autre effectue une simple ratification. Leur emploi est éminemment lié au rapport de places entre les interlocuteurs, à leur statut (ici le rapport vendeur-client) et leurs savoirs respectifs, mais également au système de préservation des faces⁸ : on peut avoir intérêt à montrer que l'on sait de quoi parle le locuteur et répondre par *voilà*, ou au contraire feindre l'ignorance et ratifier par *d'accord*. Comme on le constate, on ne se place pas dans une logique véridictionnelle, mais dans un système de rapports interpersonnels à l'intérieur duquel les locuteurs ont une certaine marge de manœuvre. Quoiqu'il en soit, étant donné la forte proportion de *voilà* et de *d'accord* dans les scènes étudiées, il serait tout à fait pertinent d'en proposer un enseignement contrasté, d'autant plus que *voilà* n'apparaît jamais comme marque d'accord dans les manuels de FLE, et que la marque d'accord, elle, est toujours uniquement présentée comme servant à répondre à une offre ou une proposition dans des échanges du type : « vous voulez faire un tennis ?/d'accord ». On peut de plus étoffer un peu le contraste entre *voilà* et *d'accord* en proposant des rapprochements avec *tout à fait* et *ok*.

3.2. *voilà/tout à fait* vs *d'accord/ok*

Ok est une marque relativement fréquente dans le corpus francophone, *tout à fait* est moins employé, mais n'apparaît jamais chez les néerlandophones. C'est une marque qui appartient, avec *absolument* ou *exactement*, à un sous-groupe de marques d'accord que l'on pourrait appeler superlatives. Or *voilà* peut permuter avec ces marques, ce qui n'est pas le cas de *d'accord* lorsqu'il est membre « évaluatif » de l'échange :

C- là vous attendez une confirmation d'Aquatour ?

⁸ Au sens de Brown et Levinson (1987) : l'image qu'un sujet met en jeu dans une interaction donnée.

E- voilà

C- d'accord ok

E- tout à fait (agence de voyage, « un père organisé »)

Dans cet échange, *tout à fait* occupe le même poste interactionnel que *voilà*, il sert à répondre à la demande de confirmation, tandis que l'ensemble *d'accord ok* constitue la ratification de la réponse jugée satisfaisante. *D'accord* peut donc s'associer avec *ok* ou permuter avec lui, comme dans l'échange suivant :

C- donc si par exemple il y avait plus de chambres on peut essayer par la Fram

E- voilà

C- d'accord

E- exactement

C- ok (agence de voyage, « un père organisé »)

A ce stade de l'analyse, on arrive donc à dessiner de nouveaux paradigmes qui rompent quelque peu avec l'effet de liste que cultivent les manuels de FLE : *voilà* et *tout à fait*, *exactement*, *absolument*, seraient à ranger, entre autres, dans la catégorie des marques d'accord-confirmation⁹, que l'on peut employer lorsqu'il y a un présupposé commun, lorsqu'on veut montrer que l'on en sait autant ou plus que son interlocuteur. En revanche, *d'accord* est avec *ok* l'expression d'une ratification plus neutre : dans le cas de l'enchaînement sur l'assertion, le locuteur considère qu'elle a bien été un « énoncé qui se présente comme ayant pour fonction principale d'apporter une information à autrui » (définition de Kerbrat-Orecchioni, 1991 : 14). L'emploi de ces marques est donc le reflet d'un rapport de faces et de savoirs présupposés. Nous terminerons par *bien sûr* parce que cette marque cristallise en quelque sorte toutes les difficultés que peut trouver le non-francophone dans l'emploi des marques d'accord, par rapport aux règles de la politesse linguistique.

3.3. *Bien sûr et la politesse linguistique*

Bien sûr est une locution adverbiale ainsi définie par le Petit Robert (1992 : 1892) : « c'est évident, cela va de soi ». Cette simple définition explique bien pourquoi avec *bien sûr* on touche à un problème de politesse linguistique et de préservation des faces : il y a en effet des situations dans lesquelles l'expression de l'évidence est impolie. Les apprenants néerlandophones ont parfois du mal à cerner les limites d'emploi de cette marque, d'autant plus qu'elle est fort appréciée des manuels de FLE qui en parsèment exagérément tous leurs dialogues. Cela aboutit à commettre des maladroites du genre suivant :

E1- je vais réserver si vous voulez ?

E2- oui bien sûr (jeu de rôle ; 6e, n°2)

Dans cet échange, on attendrait plutôt un d'accord comme réponse à une offre, *bien sûr* ici ne convient pas vraiment et constituerait même un FTA (Face Threatening Act) : il est évident que le client est venu pour réserver un séjour, et que l'employée est là pour ça, mais la politesse linguistique et la politesse tout court exigent qu'on ne lui rappelle pas son métier. En revanche, l'employée peut manifester justement une grande politesse en répondant aux demandes du client par *bien sûr* :

C- vous me faites un contrat ?

⁹ Même si leur valeur confirmative ou ratificatrice apparaît comme primordiale, il est bien entendu que ces marques n'ont pas cette unique fonction, mais l'espace imparti à cet article ne nous permet pas de faire de plus amples développements.

E- bien sûr monsieur (exemple inventé)

Les difficultés que soulève l'emploi de *bien sûr*, tantôt marque polie tantôt réponse insultante, expliquent peut-être pourquoi les francophones, même en interaction commerciale, l'utilisent avec parcimonie.

Conclusion

L'analyse de ces quelques marques d'accord, par le biais de la comparaison entre la pratique du néerlandophone et celle du francophone, permet de dégager quelques traits marquants pour l'enseignement du FLE. Elle met notamment en valeur le fait que leur emploi est fortement lié au système de place et de statut des interlocuteurs, et plus généralement à la politesse linguistique : l'évidence, par *bien sûr*, n'est pas toujours bonne à dire, et l'on n'est pas toujours habilité à confirmer par *voilà*, mais trop de *oui* ou de *d'accord* pourraient paraître peu coopératifs. La connaissance de la valeur principale de ces différentes marques permettrait ainsi au non-francophone de savoir dire *oui*, *d'accord*, *voilà*, *tout à fait* ou *bien sûr* au bon moment, mais aussi de faire l'économie d'un apprentissage fastidieux de nombreuses marques inutiles en français parlé. Dans le cadre d'un premier apprentissage des marques d'accord, il serait peut-être opportun d'enseigner les trois marques les plus fréquentes, *oui*, *d'accord* et *voilà*, pour ensuite enrichir chaque paradigme de marques différentes, même si dans tous les cas, s'engager dans la conversation comporte des risques... Encore faut-il savoir les mesurer.

Références bibliographiques

- BAYLON, Christian, *et al.* (2001). *Forum 2*, Paris : Hachette-FLE.
- BEACCO, Jean-Claude (1989). Un rendez-vous manqué ? Théorie du discours et grammaire en didactique, *Le français dans le monde, Recherches et applications* : 138-147.
- BROWN, Penelope, LEVINSON, Stephen (1987). *Politeness : some universals in language use*, Cambridge : Cambridge University Press.
- BRUXELLES, Sylvie, TRAVERSO, Véronique (2001). *Ben* : apport de la description d'un « petit mot » du discours à l'étude des polylogues, *Marges linguistiques*, 2 : 38-55.
- CHANET, Catherine (2001). 1700 occurrences de la particule *quoi* en français parlé contemporain : approche de la « distribution » et des fonctions en discours, *Marges Linguistiques*, 2 : 56-80.
- CONSEIL DE L'EUROPE (2001). *Cadre européen commun de référence pour les langues*, Paris : Didier.
- DELOMIER, Daniel (1999). *Hein* : particule désémantisée ou indice de consensualité ?, *Faits de langues*, 13 : 137-149.
- DICTIONNAIRE (1992). *Le petit Robert*, Paris.
- FLAMENT-BOISTRANCOURT, Danièle, DEBROCK, Mark (1996). Le corpus LANCOM: bilan et perspectives, *I.T.L. Review of Applied Linguistics*, 111-112 : 1-36.
- HALL, Edward Twitchell, HALL, Mildred Reed (1990). *Guide du comportement dans les affaires internationales*, Allemagne, Etats-Unis, France, Paris : Éditions du Seuil.
- KERBRAT-ORECCHIONI, Catherine (1994-1998). *Les Interactions verbales*, t.I, t.II, Paris : Armand Colin.
- KERBRAT-ORECCHIONI, Catherine (1991). *La Question*, Lyon : P.U.L.
- RIEGEL, Martin, PELLAT, Jean-Christophe, RIOUL, René (2002). *Grammaire méthodique du français*, Paris : P.U.F.

Références aux sites Internet

Département de Linguistique, Faculté des Lettres, Université de Louvain, K.U.Leuven, <http://bach.arts.kuleuven.ac.be/elicop> (2002). Consulté en juin 2006.

QUAND LES LYCÉENS ALGÉRIENS PARLENT DE LEURS COMPÉTENCES SCRIPTURALES : ENJEUX ET PERSPECTIVES DIDACTIQUES

Ali DJAROUN

alidjaroun@yahoo.fr

Laboratoire Lidilem, Université Stendhal
Grenoble

Abstract : Our article deals with Algerian high school students' representations about writing and their general lack of interest towards writing tasks at school. Paradoxically, despite this lack of interest, these students continue to think that French has to be mastered for social, professional and academic reasons. This survey shows that the heuristic value necessary for successful studies at the university is far from being salient.

1. Introduction

Le présent article expose les résultats d'une enquête réalisée en avril 2003 auprès de lycéens algériens. Cette enquête est le premier volet d'une recherche (thèse en cours) dont l'enjeu principal est de comprendre quelques-unes des raisons de l'échec massif en première année de licence de français. À titre d'illustration, ces cinq dernières années, l'échec frôle les 70%. Face à cette situation, une série d'explications tend à réduire le problème à sa seule dimension linguistique. Pour notre part, nous pensons que les difficultés peuvent être liées aussi aux conceptions que les lycéens ont de l'écriture et de son apprentissage, conceptions qui évoluent ou doivent évoluer lorsqu'ils passent du statut d'élève à celui d'étudiant. Nous posons en effet l'hypothèse que leurs représentations de l'écriture et de ses usages présentent des décalages par rapport aux représentations qu'il faut construire pour aborder avec efficacité les exercices et tâches demandés en première année de licence de français et prévus dans les programmes algériens.

L'importance que nous accordons à la contribution du sujet apprenant dans la dynamique des apprentissages nous amène à nous situer dans une perspective didactique qui pose l'hypothèse selon laquelle les représentations de l'écriture doivent être prises en compte dans les processus d'apprentissage de la production d'écrit. Comme cadre de référence, nous prenons le modèle du rapport à l'écriture, développé par un certain nombre de chercheurs (Dabène, 1987 ; Reuter, 1996 ; Barré-De Miniac, 2000).

2. Méthodologie de recueil de données

Sur la base de cette hypothèse théorique, nous avons mis en place un dispositif de recueil de données qui s'appuie sur plusieurs procédures et les croise : questionnaires, documents institutionnels (IO), analyse de sujets d'évaluation (baccalauréat, examens à la faculté) et de productions écrites (lycée, université). Dans le présent article, nous exposons quelques résultats de l'enquête menée par questionnaires, auprès de 300 lycéens de la région de Tizi-Ouzou (Algérie) appartenant à 9 classes de terminale (5 classes de section scientifique et 4 classes de section littéraire). Le questionnaire soumis comporte 16 items se référant aux dimensions du rapport à l'écriture issues du modèle développé par Barré-De Miniac (2000) : *investissement ; valeurs et attentes associées à l'écriture ; conceptions de l'écriture et de ses modalités d'acquisition*.

3. Les objectifs de l'enquête

L'enquête vise à déceler quelques facettes du rapport à l'écriture qu'ont construit des lycéens au terme de leur scolarité secondaire. Pour ceux qui s'inscrivent en première année de licence de français, ces facettes constituent un « déjà-là » potentiel au travers duquel ils appréhendent les enseignements universitaires.

4. Les résultats

Dans le sillage des représentations qu'ils se font de l'écriture, les lycéens interrogés rendent compte de l'existence d'une relative désaffection des tâches scolaires d'écriture (résumé et discussion). Cependant, cette désaffection ne les empêche pas d'assigner à la maîtrise du français des valeurs sociales et des attentes professionnelles et universitaires. Par ailleurs, ils définissent assez clairement l'écriture et ses modalités d'acquisition. Parmi les définitions proposées, la valeur heuristique, déterminante pour la réussite à l'université, est loin d'être prégnante et partagée par l'ensemble des enquêtés. Nous détaillons ces résultats dans les paragraphes qui suivent.

4.1. De l'investissement des sous-catégories de la discipline à la décision d'inscription en licence de français

Dans le but d'explorer la dimension *investissement* du rapport à l'écriture, les lycéens sont invités à se prononcer à la fois sur les types d'investissement qu'ils ont vis-à-vis des sous-catégories de la discipline de français (lecture, outils de la langue, écriture), et sur la licence de français comme projet professionnel d'avenir. Les réponses avancées donnent à voir un type d'investissement hétérogène (cf. tableau1). Les élèves disent « *Aimer bien* » la lecture, la grammaire, la conjugaison, mais pas du tout les deux tâches scolaires d'écriture (discussion, résumé). Par ailleurs, nous avons observé que cette désaffection pour l'écriture prend de très fortes proportions (65%) chez les élèves qui n'envisagent pas de poursuivre, en cas de réussite au baccalauréat, la licence de français.

4.1.1. Types d'investissement : entre hétérogénéité, désintérêt et dissonance

C'est ce qui ressort des réponses à une question demandant aux élèves d'assigner des valeurs « *Ne pas aimer* », « *Ne pas aimer du tout* », « *Aimer* », « *Aimer bien* » à la lecture, aux outils de la langue (vocabulaire, grammaire et conjugaison), et à l'écriture (résumé, discussion). La répartition globale des pourcentages attribués à ces différentes sous-catégories donne à voir un investissement hétérogène, et dissonant :

Valeurs	« Ne pas aimer » et « Ne pas aimer du tout »	« Aimer » et « Aimer bien »
Lecture	15,91%	60,57%
Vocabulaire	21,70%	48,59%
Grammaire	22,86%	54,24%
Conjugaison	25,86%	44,94%
Résumé	64,28%	19,37%
Discussion	50,40%	29,38%

Tableau 1 - Investissement des sous-catégories de la discipline de français

S'agissant de la valeur « *Aimer et aimer bien* », c'est la lecture qui enregistre le pourcentage le plus élevé avec 60%. Les deux autres catégories (outils de la langue et écriture) frôlent respectivement les 50% et 25%. Cependant, il n'en est pas de même pour les tâches scolaires d'écriture. Ce déséquilibre et l'avance enregistrée par la lecture peuvent être une conséquence de l'enseignement du français dans le secondaire qui, en Algérie, s'appuie davantage sur le texte dans une approche de réception que de production.

Le tableau montre que plus de la moitié des élèves n'aiment pas le résumé et la discussion. On peut penser que ce désintérêt tient, entre autres, à deux raisons principales. La première pourrait être d'ordre générique. Nous faisons l'hypothèse en effet que les élèves ne maîtrisent pas, voire ne disposent pas de représentations précises des règles de codification et de structuration non linguistiques des textes à produire (Rastier, 1989), en rapport avec les tâches d'écriture demandées. Cette hypothèse peut être posée dans la mesure où 44% des élèves interrogés considèrent, en réponse à une autre question (cf. 4.3.4), qu'il faut, pour écrire, avoir des modèles en tête. La deuxième relèverait de « l'image scripturale » dévalorisante qu'ils se font de leurs textes par rapport aux textes des auteurs de la littérature qu'ils étudient à l'école (Delcambre & Reuter, 2002). Se mesurant souvent à l'aune de ces derniers, ils ont construit une « image de soi » comme de piètres scripteurs. En effet, quand il s'agit d'attribuer la valeur de « véritable écrit », les écrits qu'ils produisent en milieu scolaire ne sont majoritairement pas qualifiés comme de véritables écrits, contrairement à ceux d'écrivains et de journalistes (cf. 4.3.2).

Par ailleurs, nous observons un écart de 25% entre l'investissement vis-à-vis des outils de la langue et vis-à-vis de l'écriture. Les élèves préfèrent les exercices de grammaire aux travaux d'écriture. Étant donné cet écart, on peut se demander s'il n'y a pas, dans leurs conceptions sur les modalités d'acquisition de l'écriture, une certaine dissonance entre compétence linguistique et savoir écrire.

4.1.2. Investissement vis-à-vis des tâches scolaires d'écriture et choix de la licence de français

L'enquête a été réalisée à un mois de l'examen national du baccalauréat. En cas de réussite, les élèves seront amenés à choisir parmi un certain nombre de disciplines universitaires. Dans cette perspective, nous leur avons demandé d'indiquer leurs cinq premiers choix parmi les disciplines suivantes : informatique ; médecine ; journalisme ; licence de français ; licence d'anglais ; économie ; droit ; histoire ; éducation physique ; chirurgie dentaire. Les résultats obtenus confirment l'observation empirique selon laquelle le choix de la licence de français n'est pas le choix préférentiel pour une majorité d'élèves. En effet, les premiers choix sont la médecine (67,36%), l'informatique (55,07%) et le journalisme (45,66%). La licence de français, comme premier choix, ne dépasse pas les 15%. Sur cette position, de manière inattendue (Baudelot & Establet, 1992), nous constatons que c'est le premier choix d'orientation d'un tiers des garçons de la filière Lettres. Cependant, si son pourcentage est faible en première position, la licence de français arrive avec des pourcentages non négligeables en troisième (25%) et en quatrième position (34%).

De surcroît, l'analyse transversale des réponses des élèves sur les choix disciplinaires et sur les types d'investissement vis-à-vis des tâches scolaires d'écriture indique une corrélation entre l'engouement pour la licence de français et les goûts pour l'écriture scolaire. Les élèves qui se sentent bien en production écrite (résumé, discussion) ont tendance à choisir la licence de français. À l'inverse, ceux qui se sentent « *mal* » dans ces deux tâches scolaires d'écriture boudent la filière ou bien la choisissent en quatrième ou en cinquième position.

En somme, au-delà des indications avancées sur les types d'investissement vis-à-vis des sous-catégories de la discipline de français, les élèves clarifient sur un autre registre les débats sur les conditions d'échec en première année de licence de français en Algérie.

4.2. Valeurs sociales et attentes socioprofessionnelles associées à l'écriture

Pour aller plus loin, intéressons-nous à deux questions se rapportant à la dimension *valeurs et attentes* associée à l'écriture en langue française. La première les invite à penser et à juger le cas d'un élève de terminale qui n'arrive pas à rédiger une demande d'emploi. La deuxième

leur demande de dire ce qu'ils attendent, en tant que lycéens, du savoir écrire et de la maîtrise du français en général.

4.2.1. Savoir écrire en français valorise socialement

La question relative à l'incapacité d'un élève de terminale d'écrire une demande d'emploi a suscité plusieurs opinions chez nos enquêtés. Des opinions qui montrent à quel point l'écriture est investie de très fortes valeurs sociales. En substance, comme le signalent les exemples ci-dessous, les réactions oscillent entre condamnation (cf. ex 1), dramatisation (cf. ex 2), dédramatisation (cf. ex 3) et victimisation (cf. ex 4) :

1. « Cet élève est nul, idiot, ne donne pas de l'importance et du respect pour la langue française ».
2. « Il ne mérite pas sa place en terminale, quelle est la société qui va l'accepter, il n'aura pas d'avenir ».
3. « Il ne faut pas le juger sans connaître les raisons, il n'est pas le seul, même les Français ne savent pas écrire la demande d'emploi. Ce n'est pas la fin du monde, il est là pour apprendre ».
4. « Il est innocent, victime du système éducatif barbare et irrationnel ».

Les quatre tendances dégagées confirment le prestige du savoir écrire en français dans le contexte algérien (Blanchet, 2006) et témoignent de l'importance de l'extrascolaire dans la vie des élèves. En effet, les différentes attitudes suscitées dévoilent des rapports globaux à la langue et à un de ses usages (situation de recherche d'emploi), dont ils regrettent par ailleurs l'absence dans le dispositif de formation.

4.2.2. Maîtriser le français pour s'adapter à l'université

L'adaptation aux exigences de l'enseignement supérieur est attendue de la maîtrise de la langue française par 70% des élèves interrogés. Cette attente est consciente et pragmatique car l'enseignement supérieur en Algérie, surtout pour les disciplines phares (médecine, informatique) se fait en français. La question qui mérite réflexion est de savoir si les grandes lignes du programme de français dans le secondaire prennent réellement en charge ce genre de contraintes. Cette question importante est posée dans un rapport d'experts en charge d'une mission sur les difficultés des étudiants à l'entrée à l'université (Barré-De Miniac *et al.*, 2002). Ce rapport indique que les difficultés en français entravent le suivi des études aussi bien en licence de français qu'en médecine.

Derrière les attentes exprimées à l'égard de la maîtrise du français, les lycéens soulèvent donc la question de la continuité des enseignements entre le secondaire et le supérieur en Algérie.

4.3. *L'écriture cristallise une pléthore de significations*

Dans la perspective de définition de l'acte d'écriture, les sujets interrogés proposent une pléthore de significations. Et, quand il s'agit de rendre compte des procédures de son acquisition, certains dissocient les savoirs sur la langue et les compétences en écriture.

4.3.1. L'écriture : entre « quelque chose à dire » et « quelque chose à faire »

C'est ce qui se dégage des propos de nos enquêtés interrogés sur les conceptions qu'ils se font de l'écriture et de ses modalités d'acquisition. En effet, chaque sujet est amené à construire cinq phrases de type « *écrire pour moi signifie...* ». Plusieurs aspects sont évoqués ; ils oscillent entre expression de sentiments (39%), apprentissage et maîtrise de la langue (30%), exercice de style (18%), création et construction des connaissances (15%).

Sur la question de la continuité des enseignements entre lycée et université, nous avons ici une piste à explorer. En effet, à l'instar des étudiants de langue française (Derive & Fintz, 1998 ; Rey, 1999 ; Dezutter & Thyron, 2002), les lycéens algériens n'évoquent pas la valeur heuristique de l'écriture. Pour eux, l'écriture n'est pas considérée comme un moyen de construction et de diffusion du savoir. Or, sans cette dimension, il est difficile de s'approprier les caractéristiques des discours universitaires (Pollet, 2001), et dans une large mesure, la culture universitaire (Daunay, 2002) ; une culture qui fait appel à des procédures lecturales et scripturales sous-jacentes à la conception de l'écrit telle qu'elle est envisagée dans le sillage des travaux de Goody (1977-1979). On peut se demander si cette absence a un rapport, entre autres, avec les cadres théoriques d'apprentissage du français dans le secondaire en Algérie. En effet, les Instructions Officielles ne prennent pas en charge une des voies de la conscientisation de la valeur heuristique : la dimension métacognitive.

Par ailleurs, pour l'enseignement universitaire à orientation littéraire, une proposition du type de celle d'un élève de terminale, « *Écrire pour moi, c'est comme une balle brûlante dans mon cœur frais, tel un fruit naturel, c'est l'occasion ou bien une raison de faire un signe de vie* », pourrait constituer un des appuis pour l'exploration des dimensions du littéraire, voire une des voies d'appropriation de ce discours littéraire qui pose tant de problèmes aux étudiants inscrits en première année de licence de français en Algérie.

4.3.2. L'univers des pratiques scripturales est hiérarchisé

Les élèves sont invités à dire si les écrits suivants sont de « *véritables écrits* » : une notice de médicament, un proverbe, un résumé, un exercice de grammaire, une lettre administrative, une lettre d'amour, un extrait de roman, un article de presse, un poème. Les réponses avancées font apparaître que l'univers des pratiques scripturales des élèves est stratifié en trois domaines : le scolaire, l'extrascolaire et le littéraire. Pour les élèves, les véritables écrits sont d'abord les écrits littéraires (l'extrait de roman, 62,94%), ensuite les écrits extrascolaires (la lettre d'amour, 43,60%, la lettre administrative, 37,14%) et enfin les écrits scolaires (le résumé, 19%, l'exercice de grammaire, 9,44%).

Cette stratification confirme deux choses. La première concerne le désintérêt affiché pour les tâches scolaires d'écriture, précédemment cité (cf. 4.1.1). La deuxième se rapporte à l'importance que revêt l'extrascolaire chez les élèves (cf. 4.2.1), un domaine dont l'importance pour le développement des compétences scripturales est démontrée par plusieurs recherches récentes en didactique de l'écriture (Barré-De Miniac, 2001 ; Penloup, 1999).

4.3.3. Les univers des pratiques ne sont pas complètement étanches

Si l'univers des pratiques est scindé, cette stratification n'est pas synonyme cependant d'étanchéité absolue pour les sujets interrogés. En effet, à l'instar des écrits littéraires (extrait de roman, poème), certains écrits qu'ils produisent à l'école, à savoir le résumé et l'exercice de grammaire, reçoivent également la valeur de « *véritable écrit* ». Leurs pourcentages respectifs, sur cette valeur, sont de l'ordre de 24% et de 9,44%. Certes, ils représentent un faible pourcentage, mais ils sont loin d'être négligeables. Pour un enseignement universitaire à dominante littéraire (le profil actuel de la licence de français en Algérie), il y aurait là une piste à explorer pour l'entrée dans le littéraire, dans la mesure où les élèves ne restent pas indifférents aux dimensions esthétiques de leurs gestes d'écriture, même dans les situations où il est question uniquement de l'apprentissage des règles de la grammaire. « *La tentation du littéraire* », pour reprendre l'expression de Penloup (2000), est en gestation, et son introduction dans le dispositif de formation pourrait être d'une grande aide pour les étudiants qui s'inscrivent en première année de licence de français. En effet, elle peut les aider à prendre conscience que leur geste d'écriture renferme des potentialités susceptibles de les

rapprocher des discours littéraires qu'ils n'appréhendent pas souvent avec aisance en rentrant à l'université.

4.3.4. Les savoirs sur le fonctionnement de la langue sont nécessaires mais insuffisants

Concernant les modalités d'acquisition de l'écriture, les sujets interrogés évoquent plusieurs voies. Pour la majorité, savoir écrire est une question de lecture (75%) et parfois même de la lecture d'œuvres d'écrivains français (42%), de présence d'idées préalables (48%), de maîtrise des modèles d'écriture (44%). Seuls 27 % d'élèves pensent que l'écriture dépend de la maîtrise des savoirs sur la langue. Ce pourcentage est à interroger dans la mesure où il donne à voir que 63% observent une certaine distance entre maîtrise de la langue et apprentissage de l'écriture. C'est dire que pour la majorité des lycéens, l'acte d'écriture fait intervenir des contraintes qui ne sont pas uniquement d'ordre linguistique.

Par ailleurs, les pourcentages élevés de la référence à la lecture (75%) soulèvent la question des relations ou des « interactions » entre la lecture et l'écriture. Il semblerait que la conception sous-jacente aux propos des élèves envisage ces relations ainsi : c'est la lecture qui influe sur l'écriture. On peut penser que cette orientation est le fruit des choix didactiques des enseignements, qui privilégient davantage le texte dans une approche de réception plutôt que de production. En effet, dans le contexte algérien, les activités de lecture précèdent toujours celles d'écriture. Sur ce point, les réponses des lycéens algériens ne sont pas surprenantes ; les données de certaines recherches réalisées dans d'autres contextes montrent également que les effets de la lecture sur l'écriture sont loin d'être clairs pour les élèves de différents niveaux de scolarisation (Giguère, 2002).

De surcroît, 48% des lycéens de notre enquête disent que pour écrire, « *il faut avoir des modèles en tête* ». C'est sans doute le manque de représentations claires des modèles scolaires du résumé et de la discussion qui pourrait expliquer le traitement que font les élèves des tâches d'écriture, analysé dans leurs copies.

5. Conclusion

Finalement quand les lycéens parlent de leurs pratiques scripturales, ils soulèvent des questions touchant aussi bien aux conceptions de l'écriture qu'à ses modalités d'acquisition. Des questions qui pointent autant d'enjeux cruciaux pour la continuité des enseignements entre le secondaire et le supérieur dans le contexte algérien.

Références bibliographiques

- BARRÉ-DE MINIAC, Christine (2000). *Le rapport à l'écriture. Aspects théoriques et didactiques*. Villeneuve d'Ascq: Presses Universitaires de Septentrion.
- BARRÉ-DE MINIAC, Christine (2001). L'écriture socio-professionnelle: un genre utile pour la didactique de l'écriture. *Lidil*, 23 : 187-206
- BARRÉ-DE MINIAC, Christine, BONN, Charles, BOUCHARD, Robert, COSTE, Daniel & SIBLOT, Paul (2002). Constats sur l'enseignement du français dans le supérieur en Algérie. *Mission d'étude effectuée du 17 au 21 Juin 2002*. (Rapport non publié).
- BAUDELLOT, Christian & ESTABLET, Roger (1992). *Allez les filles!* Paris : Seuil.
- BLANCHET, Philippe (2006). Le français dans l'enseignement des langues en Algérie: d'un plurilingue de fait à un plurilinguisme didactisé. *La lettre de l'AIRDF*, 38 : 31-36.
- DABÈNE, Michel (1987). *L'adulte et l'écriture. Contribution à une didactique de l'écrit en langue maternelle*. Bruxelles: De Boek-Wesmael.
- DAUNAY, Bertrand (2002). Types d'écrits et modes de production, normes, représentations, attentes institutionnelles. *Enjeux*, 53 : 213-218.

- DELCAMBRE, Isabelle & REUTER, Yves (2002). Images du scripteur et rapports à l'écriture. *Pratiques*, 113/114 : 7-28.
- DERIVE, Marie Jo & FINTZ, Claude (1998). Quelques pratiques implicites de l'écrit à l'Université. *LIDIL*, 17 : 43-56.
- DEZUTTER, Olivier & THYRION, Francine (2002). Le rapport au scriptural des étudiants en langue et littérature romanes à l'UCL: état des lieux et propositions. *Enjeux*, 53 : 139-149.
- GIGUÈRE, Jacinthe (2002). Les relations entre lecture et l'écriture : représentations d'élèves de différents niveaux d'habileté. *Les cahiers Théodile*, 2 : 39-51.
- GOODY, Jack. (1979). *La raison graphique. La domestication de la pensée sauvage*. Paris : Minuit.
- PENLOUP, Marie-Claude (1999). *L'écriture extrascolaire des collégiens. Des constats aux perspectives didactiques*. Paris : ESF éditeur.
- PENLOUP, Marie-Claude (2000). *La tentation du littéraire. Essai sur le rapport à l'écriture littéraire du scripteur ordinaire*. Paris : Didier- Credif.
- POLLET, Marie-Christine (2001). *Pour une didactique des discours universitaires. Étudiants et système de communication à l'université*. Bruxelles : De Boeck Université.
- RASTIER, François (1989). *Sens et textualité*. Paris : Hachette.
- REUTER, Yves (1995). Les relations et les interactions lecture-écriture dans le champ didactique. *Pratiques*, 86 : 5-23.
- REUTER, Yves (1996). *Enseigner et apprendre à écrire*. Paris : ESF éditeur.
- REY, Bernard (1999). Les compétences exigées à l'université : quelques hypothèses récentes, in *Enseignement secondaire et enseignement universitaire. Quelles missions pour chacun ?*, Louryan, Stéphane & Thys-Clément, Françoise (Eds). Bruxelles : Éditions de l'Université de Bruxelles, 57-67.

ÉTUDE ET OPÉRATIONNALISATION DE PHÉNOMÈNES DE STÉRÉOTYPIE POUR LA DIDACTIQUE DE L'ÉCRITURE

Bernadette KERVYN¹

bkervyn@wanadoo.fr

IUFM d'Aquitaine

47000 Agen

France

Cedill (Centre de recherche en didactique des langues et littératures romanes)

1348 Louvain-la-Neuve

Belgique

Abstract : This paper seeks to answer two questions which concern the teaching-learning process of writing in the last junior classes of the French primary school. Firstly: why is it important to examine language forms of stereotypy phenomena ? And secondly: how can validity teaching applications be developed? The aim is to guide the creation of such applications by studying language forms of stereotypy phenomena as well as their operationalisation.

Quels intérêts peut présenter la prise en compte de phénomènes de stéréotypie, sous leur forme langagière, pour l'enseignement-apprentissage de l'écriture en fin de cycle 3 de l'école élémentaire² ? Comment construire la validité de cet outil potentiel ? Nous tenterons ci-dessous d'esquisser des réponses à ces deux questions ancrées dans la didactique du français.

1. Cadre et objectif de notre travail

Les éléments de réponse qui vont suivre, d'ordre à la fois épistémologique, méthodologique et praxéologique, font partie intégrante de notre recherche doctorale visant à examiner et à construire, théoriquement et pratiquement, la pertinence de phénomènes de stéréotypie pour l'enseignement-apprentissage de l'écriture à l'école élémentaire. Dans ce projet, nous validons théoriquement la pertinence de l'outil en prenant appui sur des auteurs de référence³ et nous soumettons cette validité théorique au test des pratiques, par une démarche praxéologique qualitative de construction et d'utilisation d'outil(s) didactique(s) dont il sera ici principalement question.

Par ce double mouvement, notre recherche se situe dans un des rôles possibles de la didactique dont parlent Bronckart et Plazaola Giger : « À nos yeux, le champ didactique constitue le lieu majeur de *validation externe*⁴ de nombre de propositions théoriques émanant des sciences sociales/humaines (...). La mise en circulation de ces propositions dans les pratiques didactiques constitue un des tests possibles de leur validité intrinsèque. » (1998 : 56). De même, ces choix de recherche laissent transparaître une vision praxéologique de la didactique du français et des recherches en didactique, où les pôles des savoirs, des élèves et des enseignants se définissent mutuellement, dans leur relation aux autres (Halté, 1998), et

¹ Article rédigé en collaboration avec Jérôme Faux, école d'application Bara, 47000 Agen.

² C'est-à-dire pour des élèves âgés majoritairement de 10-11 ans, situés, dans le système français, en classe de CM2 (Cours Moyen 2).

³ Bakhtine (1984), Bernié (2001), Chabanne et Bucheton (2002), Clot (1999), Garcia-Debanc (1990)...

⁴ En italiques dans le texte.

amènent à traiter des objets d'enseignement-apprentissage dans leur complexité et dans leur lien potentiel aux travaux d'autres disciplines alors contributoires.

C'est dans ce cadre de traitement des objets d'enseignement-apprentissage, de références académiques, de proximité du terrain que prend place notre étude des phénomènes de stéréotypie dans leur forme langagière ainsi que dans leur opérationnalisation⁵ dont nous limiterons ici la présentation à un exemple qui sera analysé davantage du point de vue de l'enseignant.

2. Dispositif d'opérationnalisation

Pour construire la pertinence des phénomènes de stéréotypie comme outil didactique, nous avons mené, durant l'année 2004-2005, une recherche-action impliquant 8 personnes (7 enseignants collaborateurs et nous-même) en situation d'hétérogénéité fonctionnelle⁶ ainsi que 2 classes de CM2. Cette recherche-action incluait un temps de formation professionnalisante de 5 séances au terme desquelles nous avons filmé la manière dont 2 enseignants collaborateurs de CM2 travaillaient l'écriture littéraire puis non littéraire avec leurs élèves, en s'appuyant sur des phénomènes de stéréotypie de leur choix. Malgré la prise de risque importante, nous avons souhaité laisser ces professeurs assez libres dans leur élaboration d'outil(s) didactique(s) à partir des approches théoriques et pratiques envisagées lors du temps de formation et présentées ci-après. Cette liberté, dans un cadre concerté, manifeste autant notre reconnaissance de leurs compétences d'enseignant chevronné que notre volonté d'adopter une démarche écologique, en nous situant au plus proche des réalités du terrain. Néanmoins, pour garantir une recherche de type qualitatif, nous avons veillé au recrutement d'enseignants fortement qualifiés (finalement, pour différentes raisons dont une nécessaire disponibilité horaire, tous maîtres formateurs) et ouverts à l'innovation professionnelle (Cros, 2001), à l'appropriation préalable par les enseignants des phénomènes de stéréotypie et de travaux théoriques et didactiques sur le sujet, à une réflexion commune sur un maximum de pistes didactiques possibles avant la confection des séquences par chacun des 2 enseignants expérimentateurs.

3. Phénomènes de stéréotypie et travaux envisagés

Qu'entendons-nous par « phénomènes de stéréotypie » ?⁷ A la manière d'Amossy et d'Herschberg Pierrot (1997), nous utilisons cette étiquette pour désigner tant des produits (stéréotype, cliché, lieu commun, prototype, idée reçue, topoï) que des processus (stéréotypisation, stéréotypage, clichage) étudiés en sciences humaines et plus précisément en sciences sociales, en littérature et en sciences du langage.

Nous privilégions cette étiquette par rapport à celle plus courante de « stéréotype » ou de « stéréotypage » pour rendre apparente la pluralité des recherches menées et des notions considérées, ainsi que pour marquer une distance par rapport à l'acception péjorative communément attribuée au terme « stéréotype ».

Dans le cadre de notre objectif d'opérationnalisation didactique, nous avons choisi de conserver au maximum cette diversité lors du travail initial de formation avec les enseignants

⁵ Par opérationnalisation, nous entendons le passage ou la transformation d'un savoir principalement théorique en un savoir stratégique pour l'action d'enseignement-apprentissage en contexte scolaire (Van den Maren, 2003).

⁶ Nous avons au sein du groupe des rôles différents : professeur expérimentateur, professeur ressource, garant de la démarche et du savoir savant. Quel que soit le rôle et la plus ou moins grande implication de ces collaborateurs, nous tenons à les citer en guise de remerciement : Nadine Boissié, Jérôme Faux, Thierry Lamarque, Nathalie Morales, Françoise Roland, Maïté Rouire, Viviane Trignac.

⁷ La notion étant peu usitée et non stabilisée, il nous semble important d'en tracer les contours.

collaborateurs car tous ces phénomènes de stéréotypie nous semblaient potentiellement intéressants pour la didactique de l'écriture. Dans ce qui suit, puisque nous avons choisi de nous focaliser sur l'opérationnalisation de ces phénomènes de stéréotypie, nous ne ferons que citer les produits et processus de stéréotypie envisagés avec les enseignants collaborateurs⁸ en vue de transformer certains de ces phénomènes de stéréotypie en outil(s).

Par rapport aux études pointues sur ces phénomènes, il est évident que notre travail collectif ne visait pas l'approfondissement mais a consisté en un survol suffisamment riche pour fournir aux enseignants collaborateurs des bases leur permettant de faire de phénomènes de stéréotypie un ou des outils d'écriture pour leur classe.

Du côté des sciences sociales, il a été question surtout du stéréotype entendu comme une représentation mentale culturelle préconçue, produit de l'imaginaire social, non fondé sur l'expérience et par lequel nous typifions ou généralisons le réel. Bivalente, cette manière de penser est très souvent dénoncée pour son caractère réducteur et réhabilitée par certains auteurs pour sa participation à l'identité sociale (Fischer, 1996). Le stéréotype est alors envisagé comme un facteur de cohésion sociale. La fonction constructive du stéréotype est aussi liée au processus cognitif incontournable de catégorisation et de généralisation. Nommé « stéréotypage » ou « stéréotypisation » selon les chercheurs, il consiste à appliquer à des individus « un jugement stéréotypique qui rend ces individus interchangeables avec les membres de leur catégorie » (Leyens, Yzerbyt et Schadrin, 1996 : 24).

Du côté de la littérature, avec le terme « cliché » désignant une formule de style figée, une expression usée, opposée à la création personnelle ou signe de culture de masse, ce sont les questions de l'originalité et de l'esthétisme qui ont primé. Ici encore nous avons relevé que cette forme langagière surtout dénigrée est revalorisée, dans certaines études, pour son efficacité textuelle, pour sa fonction dans la construction de l'intertextualité (Riffaterre, 1979), pour sa participation à l'inscription de l'histoire et de la société dans les textes (Amossy, 1991), ou encore parce que la langue se distingue du discours singulier où l'auteur énonce sa vérité (Paulhan, 1967). En lecture littéraire, les stéréotypes sont également appréciés pour leur valeur intertextuelle, générique et culturelle (Eco, 1985). Lire relève en partie d'une projection de stéréotypes, c'est-à-dire d'un processus de « clichage » (Dufays, 1984) ou de stéréotypisation.

Du côté des sciences du langage, nous pouvons signaler le travail de Gross (1996) sur des locutions figées, des clichés et des expressions figées selon des critères syntaxiques et sémantiques, celui de Putnam (1985) sur le stéréotype envisagé comme une partie de la signification liée à une représentation simplifiée, conventionnelle et courante, et qui permet une communication optimale en société dans la mesure où il véhicule la norme sociale et culturelle. Dans une même logique sémantique mais aussi psycholinguistique de catégorisation cognitive, nous avons considéré la notion de prototype définie comme « le meilleur représentant de sa catégorie » ou comme des « combinaisons de propriétés typiques » (Kleiber, 2004 : 185-186). Parallèlement, il existe des approches plus pragmatiques consacrées aux lieux communs, aux topoï ou idées reçues, collectivement admises, polyphoniques et qui servent à asseoir tout discours persuasif (Eggs, 1994). Enfin, il faut évoquer le terrain propice que constitue l'analyse du discours pour l'étude de la stéréotypie, les contraintes génériques et institutionnelles gouvernant le discours (Maingueneau, 1996), les mots dont le sens n'est pas séparable des contextes où ils s'insèrent

⁸ Parce qu'il a l'avantage de mentionner de façon claire et structurée un nombre important d'études sur le sujet, nous avons pris comme base principale l'ouvrage *Stéréotypes et clichés* (Amossy et Herschberg Pierrot, 1997) pour notre travail avec les enseignants collaborateurs.

ni de la place des locuteurs dans le champ sociohistorique et institutionnel (Amossy et Herschberg Pierrot, 1997).

Du côté de la didactique, la majorité des travaux consultés concerne la lecture littéraire. Les activités portent généralement sur le repérage de clichés ou de stéréotypes et parfois sur leur traitement par l'auteur (participation, mise à distance par la critique, la parodie...) (Dufays, Gemenne, Ledur, 1996 ; revue *Recherches*, 1989). Elles sont fréquemment utilisées pour découvrir ou identifier un genre ou un type de texte et sont parfois prolongées par une tâche de production écrite où la bivalence des stéréotypes et des clichés semble peu considérée. De même, les pratiques de mise en réseau recommandées à l'école primaire par le Ministère de l'Education Nationale (2002) se basent parfois aussi sur de la stéréotypie, rarement identifiée en tant que telle. Il en va de même pour l'écriture où des modèles sont souvent requis mais sans distinction claire entre modèle prototypique, règle(s) du genre et modèle singulier provenant de la lecture d'un ou de plusieurs auteurs. Enfin, ajoutons qu'en lecture comme en écriture, des auteurs présentent la découverte de connaissances stéréotypées comme un premier accès important à la culture attendue par l'école (Rouxel, 1997, Dufays et Kervyn, 2003).

De ce bref parcours, il ressort que les études sur les phénomènes de stéréotypie envisagés concernent aussi bien des représentations, des images mentales que des éléments textuels ou discursifs, des produits que des processus, que ceux-ci sont tantôt envisagés péjorativement tantôt de façon constructive mais souvent comme un élément incontournable de l'interaction sociale.

Assez rapidement, le potentiel de ces divers travaux pour la didactique de l'écriture et pour l'enseignement-apprentissage de l'écriture dans nos contextes professionnels a émergé. Parmi les intérêts relevés par les enseignants collaborateurs, l'on trouve : la possibilité de partir du déjà-là chez les élèves, l'ancrage ou l'insertion dans de la norme sociale, la distanciation par rapport à la norme pour créer du différent, l'opération de réduction à de l'essentiel, l'aspect incontournable de ces phénomènes et la nécessité de les rendre conscients pour qu'ils soient moins dangereux et qu'ils ne bloquent pas l'apprentissage.

Concrètement, au terme de ce temps de formation-échange portant sur ces travaux, sur leur intérêt et leurs prolongements didactiques possibles, nous avons récolté 4 séquences portant sur l'écriture scolaire via le traitement de phénomènes de stéréotypie (deux séquences sur l'écriture poétique, une sur l'écriture du compte rendu scientifique et une dernière sur l'écriture de pages documentaires), les productions des élèves (productions initiales, intermédiaires et finales), leur réponse à des questionnaires en début et fin de parcours ainsi que l'entretien avec certains d'entre eux. Ces données relativement substantielles n'étant pas encore toutes analysées, nous nous limiterons dans le présent article à exposer la séquence concernant l'écriture du compte rendu scientifique⁹ – réalisée par Jérôme Faux – ainsi qu'une partie de la réflexion qui l'a entourée. Faisant partie du processus de recherche, notre objet n'était pas et n'est pas d'obtenir une séquence hors de toute critique ou au label séducteur mais contestable de « modèle à reproduire ». Il s'agit, comme il a déjà été dit, de construire l'outil « phénomènes de stéréotypie ».

⁹ Une des deux séquences concernant l'écriture poétique vient de faire l'objet d'un article (à paraître début 2007 dans les actes du colloque international « Stéréotypage, stéréotypes : fonctionnements ordinaires et mises en scène » qui s'est tenu à Montpellier en juin 2006) dont certains éléments du cadrage recourent d'ailleurs ceux présentés dans ce texte.

4. Un exemple de séquence didactique intégrant le travail sur les phénomènes de stéréotypie

4.1. Pertinence d'un travail sur l'écriture du compte rendu d'expérience scientifique

Pour atteindre une maîtrise transversale de la langue, les Programmes Officiels de 2002 prévoient qu'elle soit travaillée dans chaque discipline, au cycle 3, où elle se décline sous la forme de compétences organisées autour des rubriques « Parler », « Lire » et « Écrire ». Dans les compétences attendues en fin de cycle 3, au niveau de l'écriture en sciences, figure la rédaction, avec l'aide du maître, d'un compte rendu d'expérience ou d'observation dont il est précisé qu'il s'agit d'un « texte à statut scientifique » (MEN, 2002 : 223). Globalement, dans ces nouveaux programmes, le renforcement de la maîtrise du langage et de la langue française est posé comme essentiel dans l'enseignement, entre autres, des sciences et de la technologie. Autrement dit, en sciences, on construit des compétences scientifiques et on développe des compétences langagières orales et écrites.

Ainsi, conformément aux programmes en vigueur, lors d'une séance de sciences, les élèves sont bien souvent invités à écrire. En amont d'une expérience, ils consignent, par exemple, ce qu'ils savent du sujet observé. Lors de la réalisation de l'expérience, ils prennent des notes, ils réalisent des schémas. En aval, ils expliquent par écrit ce qu'ils ont appris de l'expérience. Sous cet angle, l'écriture du compte rendu ici envisagé est donc l'une des nombreuses applications de la politique éducative concernant la maîtrise de la langue, politique répondant à l'enjeu fondamental d'intégration sociale.

De façon plus disciplinaire, on peut aussi s'interroger sur ce qu'apporte l'écriture de comptes rendus au développement chez l'élève de compétences scientifiques. Aujourd'hui, l'importance de l'écriture en sciences est communément admise¹⁰ :

« La trace écrite est la mémoire du travail de l'enfant, l'image de l'évolution de la construction de sa pensée, des progrès effectués. Elle aide à l'apprentissage, à l'expression des connaissances, à la mémorisation, à la structuration et à la conscientisation, facilite la communication, l'échange, la confrontation, l'argumentation, devient support de l'imaginaire. Ainsi tout au long de la démarche scientifique, il est indispensable de convoquer l'écrit et d'apprendre aux enfants à s'y référer, à s'y appuyer pour construire leur pensée et leur savoir. » (Béal, Lacour, Maïaux, 2004 : 109)

Par l'écriture du compte rendu, ils vont s'essayer à une activité et à un geste professionnel propres à la communauté des scientifiques, certes transposés dans une forme scolaire. En effet, le rôle d'un scientifique ne s'arrête pas à la réalisation de la recherche, il lui faut la communiquer. Dans le préface de son ouvrage, Day exprime qu'« il n'est pas nécessaire que le plombier écrive au sujet des tuyaux qu'il répare ; ni que l'avocat écrive sur ses plaidoiries (sauf peut être des petits textes ponctuels), mais le scientifique – cas sans doute unique parmi les métiers et professions – doit fournir un document écrit montrant ce qu'il a fait, pourquoi il l'a fait, comment il l'a fait et quels enseignements il en a tiré » (Day, 1989 : 11). Ainsi le scientifique ne doit-il pas seulement faire la science, mais écrire la science et rendre des comptes. Dans cette optique, pratiquer le compte rendu d'expérience en sciences, c'est lier le développement de la maîtrise de la langue à l'un de ses usages socioprofessionnels.

¹⁰ Nous renvoyons à ce sujet au PRESTE (Plan de Rénovation de l'Enseignement des Sciences et de la Technologie à l'Ecole) de 2000 ainsi qu'au site internet de La main à la pâte (<http://www.inrp.fr/lamap/>).

4.2. Présentation de la séquence

La séquence dont nous rendons compte a été élaborée au mois de mai pour une classe de CM2 présentant un profil pédagogique et social relativement homogène. Les élèves y rédigeaient depuis le début de l'année des comptes rendus d'expérience scientifique au rythme de 2 voire 3 par mois et y prenaient du plaisir.

Dans ce contexte, parmi les phénomènes de stéréotypie considérés initialement, l'enseignant collaborateur a retenu le(s) stéréotype(s) potentiellement présent(s) dans les représentations initiales de ses élèves et a choisi en conséquence de travailler sur les phases de la lune dont l'explication relève fréquemment du stéréotype. Ce choix de l'enseignant pose des questions tout à fait pertinentes pour la didactique de l'écriture et la maîtrise de la langue dans les disciplines : quelles incidences sur l'écriture du compte rendu peut avoir une représentation initiale stéréotypée du sujet étudié ? Un travail explicite sur les stéréotypes aide-t-il les élèves à rédiger leurs comptes rendus ?

Ce travail a été mené en 3 temps répartis sur 4 séances d'environ 55 minutes. Partant de l'observation de photographies de la lune prises au cours d'une lunaison, le temps 1 s'est ouvert sur le questionnement libre des élèves du phénomène et a abouti à la sélection d'une question alors reformulée par écrit. A suivi une discussion dont l'objectif était de faire ressortir les représentations initiales des élèves, également recueillies par le biais d'un questionnaire. Cette double procédure avait pour objet de conforter ou non la présence de stéréotype(s). Dans le temps 2, les élèves ont imaginé des expériences qu'ils ont schématisées. Ils ont listé le matériel nécessaire, anticipé les résultats avant de réaliser leurs expériences, d'observer et de partager leurs résultats. La confrontation des résultats, dans le temps 3, a fait apparaître la nécessité d'une nouvelle expérience, proposée cette fois par l'enseignant. Celle-ci a permis d'expliquer les phases de la lune. Le travail s'est achevé sur l'écriture individuelle du compte rendu de ces expériences (celle imaginée par l'élève et celle proposée par l'enseignant) et sur la réponse à un questionnaire concernant l'écriture du compte rendu et l'explication du phénomène lunaire.

Le schéma suivant présente de façon synthétique les différents temps de la séquence.

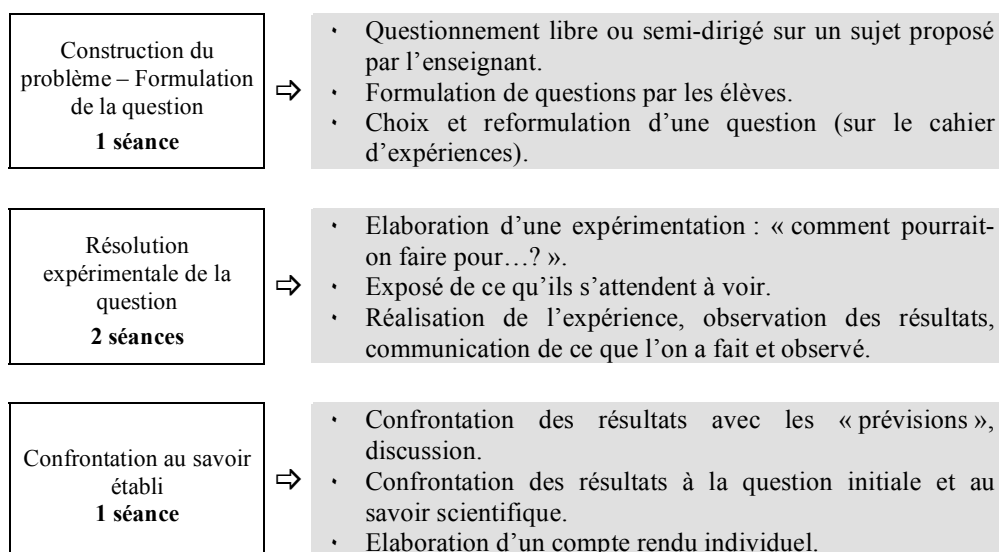


Figure 1 - Schéma de la séquence

4.3. *Analyse de la séquence en guise de conclusion*

En début de séquence, comme supposé par l'enseignant, tous les élèves de la classe partageaient la même représentation concernant les phases de la lune : l'ombre de la terre portée sur la lune expliquerait ses différentes phases (« quand les rayons du soleil tape sur la Terre, c'est l'ombre de la Terre qui va sur la Lune », note, par exemple, un des élèves). Par ses côtés réducteurs, schématisant, résistant au changement, partagé par un grand nombre et ne relevant pas de l'observation, cette représentation répondait bien aux critères de stéréotypie. Selon nous, les caractéristiques du phénomène étudié mais aussi la prégnance du stéréotype (très probablement d'ailleurs liée à la complexité cognitive à laquelle l'enfant est confronté pour expliquer le phénomène¹¹) ont fait que les expériences conçues et réalisées par les élèves n'ont pas abouti à une explication claire. Cette hypothèse, à tester, repose sur la difficulté constatée chez les enfants, malgré l'étayage du maître, à observer du non attendu et à concevoir une expérience appropriée, deux compétences qui nécessitent une mise à distance préalable, forcément difficile en cas de stéréotypie par définition prégnante et résistante au changement. Leur perplexité face à ce non-aboutissement ainsi que l'expérience proposée par l'enseignant semblent avoir mis une première fois à mal le modèle stéréotypé. Le passage à l'écriture personnelle du compte rendu ainsi que l'étiquetage explicite de leur représentation initiale comme stéréotype les ont amenés à dépasser une seconde fois cette représentation initiale au profit d'une explication scientifique plus complexe. L'on voit ainsi le rôle qu'a joué l'ensemble du dispositif dans l'appropriation progressive d'un savoir, moins stéréotypé et plus scientifique, même si ce passage est loin d'être évident. En effet, dans le questionnaire rempli en fin de séquence, beaucoup d'élèves écrivaient qu'ils préféreraient leur représentation initiale, plus « facile » et/ou « logique », au discours scientifique sur le phénomène, et ce juste après avoir correctement expliqué le phénomène.

Mais en quoi, dans cette séquence, le stéréotype est-il un outil pour l'enseignement-apprentissage de l'écriture ? L'écriture du compte rendu, dans la difficulté qu'elle a posée aux élèves, porte la trace du stéréotype initial et du mouvement cognitif en cours : ce n'est ni la mise en œuvre du protocole ni sa description qui leur a posé problème mais l'écriture des observations et l'interprétation qui en découle. La prise de conscience par l'enseignant et par l'élève d'une représentation initiale stéréotypée, dans la mesure où cette dernière est susceptible d'entraver l'écriture du compte rendu (choix des mots ou pertinence du propos), peut amener à une vigilance particulière lors de l'écriture ou à un étayage plus important du maître. Le repérage du stéréotype peut aussi apporter à l'enseignant une clé pour comprendre et évaluer la production de l'enfant. Jérôme Faux, l'enseignant collaborateur, estime que le travail sur les phénomènes de stéréotypie lui a permis d'affiner sa compréhension de la difficulté éprouvée par les élèves dans cette situation d'apprentissage. Il souligne également avoir réalisé l'écart complexe et délicat entre le savoir énoncé correctement lors de l'évaluation, le savoir auquel tient l'enfant et le savoir réellement approprié. Il dit avoir pris conscience de l'impact fortement contraignant que peut avoir une représentation initiale stéréotypée sur le déroulement d'une séquence (au scénario habituellement approprié) et de la nécessité de l'ajuster pour prendre en compte la résistance des élèves et la mise à distance problématique pour eux. En cela, notamment, il considère que le traitement des stéréotypes est un outil d'enseignement intéressant qui lui permet de faire évoluer sa pratique.

Par ailleurs, côté élèves, nous pensons, et c'est à vérifier, que le fait d'avoir explicitement situé leur représentation initiale du côté de la stéréotypie, en la posant comme problématique

¹¹ La complexité cognitive provient du décalage avec le quotidien. En effet, au quotidien nous baignons dans la lumière où un objet est non visible quand il y a obstacle (la terre dans l'explication des enfants). Par contre, si nous voyons les objets célestes, c'est parce qu'ils sont éclairés. L'explication à construire est donc fortement éloignée du quotidien.

pour acquérir de nouveaux savoirs en sciences, a facilité, pour certains, le passage à un discours plus scientifique. Évoquant l'explication de départ dans le questionnaire final, un élève écrit : « C'est un stéréotype. Il est faux mais on croyait tous qu'il était vrai ». Un autre, lors de l'entretien oral nous dit : « en sciences, le stéréotype, ça nous empêche de connaître la vérité (...) alors, c'est plus difficile, pour être scientifique, faut pas y croire ».

Références

- AMOSSY, Ruth (1991). *Les idées reçues. Sémiologie du stéréotype*. Paris : Nathan.
- AMOSSY, Ruth & HERSCHBERG-PIERROT, Anne (1997). *Stéréotypes et cliché*. Paris : Nathan.
- BAKHITINE, Mikhail (1984). *Esthétique de la création verbale*. Paris : Gallimard.
- BÉAL, Yves, LACOUR, Martine & MAÏAUX, Frédérique (2004). *Ecrire en toutes disciplines. De l'apprentissage à la création*. Paris : Bordas.
- BERNIÉ, Jean-Paul (2001). *Apprentissage, développement et significations. Hommages à Michel Brossard*. Bordeaux : Presses Universitaires de Bordeaux.
- BRONCKART, Jean-Paul & PLAZAOLA GIGER, Itziar (1998). La transposition didactique. Histoire et perspectives d'une problématique fondatrice, *Pratiques*, 97-98 : 35-58.
- CHABANNE, Jean-Charles & BUCHETON, Dominique (2002). *Parler et écrire pour penser, apprendre et se construire : l'écrit et l'oral réflexifs*. Paris : PUF.
- CLOT, Yves (Dir) (1999). *Avec Vygotski*. Paris : La dispute.
- CROS, Françoise (2001). *L'innovation scolaire*. Paris : INRP.
- DAY, Robert A. (1989, 3e éd.). *How to Write and Publish a Scientific Paper*. Cambridge : Cambridge University Press.
- DUFAYS, Jean-Louis (1984). *Stéréotype et lecture*. Liège : Mardaga.
- DUFAYS, Jean-Louis, GEMENNE, Louis & LEDUR, Dominique (1996). *Pour une lecture littéraire*. Bruxelles : De Boeck.
- DUFAYS, Jean-Louis & KERVYN, Bernadette (2003). Quels usages du stéréotype dans l'écriture au secondaire ? Observations et propositions sur base d'une centaine de copies d'élèves, *Pratiques*, 117-118 : 208-218.
- ECO, Umberto (1985). *Lector in fabula ou la Coopération interprétative dans les textes narratifs*. Paris : Grasset.
- EGGS, Ekkehard (1994). *Grammaire du discours argumentatif*. Paris : Editions Kimé.
- FISCHER, Gustave (1996). *Les concepts fondamentaux de la psychologie sociale*. Paris : Dunod.
- GARCIA-DEBANC, Claudine (1990). *L'élève et la production d'écrits*. Metz : Université de Metz.
- GROSS, Gaston (1996). *Les expressions figées en français*. Paris : Ophrys.
- HALTÉ, Jean-François (1998, 3e éd.). *La didactique du français*. Paris : PUF.
- KLEIBER, Georges (2004) (rééd.). *La sémantique du prototype*. Paris : PUF.
- LEYENS, Jean-Philippe, YZERBYT, Vincent & SCHADRON, Georges (1996). *Stéréotypes et cognition sociale*. Liège : Mardaga.
- MAINGUENEAU, Dominique (1996). *Les termes clés de l'analyse du discours*. Paris : Seuil.
- MINISTÈRE DE L'ÉDUCATION NATIONALE (2002). *Qu'apprend-on à l'école élémentaire ?* Paris : CNDP/XO Editions.
- PAULHAN, Jean (1967). *Œuvres complètes*. Cercle du livre précieux.
- PUTNAM, Hilary (1985). Signification, référence et stéréotype, *Philosophie* 5.
- RECHERCHES (1989). *Apprentis-sages et stéréotypes*. 10. Lille : AFEF.
- RIFFATERRE, Michael (1979). *La production du texte*. Paris : Éditions du Seuil.
- ROUXEL, Annie (1997). *Enseigner la lecture littéraire*. Rennes : Presses Universitaires de Rennes.
- VAN DEN MAREN, Jean-Marie (2003). *La recherche appliquée en pédagogie. Des modèles pour l'enseignement*. Bruxelles : De Boeck.

OUTILS INFORMATIQUES POUR LA LINGUISTIQUE ET LA DIDACTIQUE DES LANGUES

ESQUISSE D'UNE FORMALISATION DES ACTES DE LANGAGE INDIRECTS

Sophie ANQUETIL

sophie.anquetil@unicaen.fr

Laboratoire CRISCO

Université de Caen Basse-Normandie

14032 Caen Cedex

Abstract : This research poses the problem of the automatic treatment of indirect speech acts. How is it possible to detect these statements automatically whereas their literal meaning is different from their illocutionary meaning ? The problem is made more complicated by the fact that their statements are characterized by formal marks which do not seem to correspond to illocutionary acts.

1. Introduction

Depuis quelques années, l'essor des nouvelles technologies informatiques a permis d'ouvrir la voie à de nouveaux domaines de recherche en linguistique. Cette convergence des deux disciplines – informatique et linguistique – répondait à un besoin réciproque et croissant. En effet, la nécessité pour les technologies de faire appel à la linguistique était grandissante, tant la communication homme-machine s'avérait problématique. À l'inverse, la linguistique a pu profiter de nouveaux outils d'analyse et de bases de données quantitativement incomparables à ce que l'on aurait pu obtenir manuellement. L'intérêt actuel porté à la linguistique de corpus en témoigne. Si le recours à ce type d'outils semble aujourd'hui être la clé d'une analyse réussie, la constitution d'un corpus à l'aide de techniques informatiques est soumise à des critères de sélection et d'organisation qui déterminent un type d'étude spécifique. En effet, pour faire la recherche informatique d'un phénomène discursif, il faut au préalable en connaître les manifestations linguistiques. Et cela se révèle difficile lorsque les énoncés qui font l'objet de cette recherche ne possèdent pas de marques formelles spécifiques. C'est notamment le cas des énoncés dont le sens diverge de la forme : leur traitement est d'ailleurs l'un des plus délicats à effectuer pour un ordinateur.

2. Problématique

C'est face à ce problème que nous nous trouvons lorsque nous cherchons à effectuer un repérage automatique d'énoncés supports d'actes de langage indirects (dorénavant, ALI) : comment est-il possible de repérer automatiquement ces énoncés, sachant que leur sens littéral diverge du sens actualisé en contexte ? Par exemple, l'énoncé 1 est sur le plan illocutoire une requête (acte directif selon la terminologie de Searle, 1982) ; alors que le sens littéral de cet énoncé interroge l'interlocuteur sur ses disponibilités horaires :

1. Peux-tu passer demain après le déjeuner ?

Une autre difficulté vient se greffer à la première : la plupart des marques formelles récurrentes dans ces énoncés (les marqueurs de dérivation illocutoire) sont identiques quelle que soit leur valeur illocutoire. D'ailleurs, si nous opérons une recherche d'énoncés supports d'ALI dans Frantext, à partir de critères de sélection tels que la forme fléchie du verbe *pouvoir*¹ et la seconde personne, nous obtenons des énoncés à valeur illocutoire variable :

2. **Pourrais-tu** le retrouver ? (directif)

¹Ce verbe est une marque non décidable (Cf. 3.1. et 5.2.).

3. Tu **aurais** quand même **pu** y penser. (expressif)
4. Comment **peux**-tu dire cela ? (expressif)
5. Tu **peux** compter sur moi. (promissif)

Nous nous trouvons donc face à une difficulté méthodologique : l'élaboration d'un corpus d'étude d'ALI à l'aide d'outils informatiques.

3. Hypothèses

3.1. *Hypothèse 1 : Des marques non décidables ?*

La première hypothèse que nous émettons, part d'une recherche antérieure (Anquetil, à paraître) : les marqueurs de dérivation illocutoire (*pouvoir, falloir, devoir...*) ne sont pas des indices ou des traces linguistiques de la valeur illocutoire des ALI, mais des manifestations de l'arrière-plan représentationnel (arrière-plan qui est, notons-le, constitué d'inférences stéréotypiques). C'est pourquoi ces marques ne permettent pas de « détecter » telle valeur illocutoire spécifique : ce sont des **marques non décidables**. La mise en évidence de ces marqueurs récurrents dans les énoncés supports d'ALI permettrait donc d'aplanir une des difficultés théoriques préalablement posées, à savoir pourquoi certaines marques récurrentes ne sont pas distinctives d'une valeur illocutoire particulière.

3.2. *Hypothèse 2 : Des marques décidables ?*

La première hypothèse, même confirmée, ne permet pas à elle seule de résoudre le problème que pose l'élaboration d'un corpus d'énoncés supports d'ALI, à partir d'outils informatiques. Une seconde hypothèse vient donc compléter la première : seul l'ancrage des inférences constitutives de l'arrière-plan représentationnel, dans une situation d'énonciation, permet d'interpréter ces énoncés comme ALI. Par conséquent, les marques d'énonciation (à savoir les pronoms personnels, le temps et le mode du verbe de la principale, etc.) parce qu'elles sont autant de points d'ancrage de ces inférences dans l'énonciation, déterminent la valeur illocutoire de ces énoncés. Ce sont des **marques décidables**. Ainsi, seule une combinaison de marques décidables et de marques non décidables permet de détecter automatiquement des ALI à valeur illocutoire spécifique.

4. Le choix du cadre théorique

Les différentes approches adoptées par les linguistes et les chercheurs en TAL pour formaliser les ALI n'ont pas permis de réponse réellement satisfaisante. Nous avons donc dû élaborer notre propre cadre théorique. Celui-ci prend appui sur la théorie des topoï (Ducrot, 1988). Avant de présenter ce cadre et de confirmer les hypothèses préalablement posées, nous nous devons d'expliquer brièvement les raisons qui ont motivé ce choix.

4.1. *L'approche intentionnelle : une approche problématique*

Le problème que pose toute approche intentionnelle lorsque l'on cherche à formaliser les actes de langage et les structures intentionnelles est l'amalgame entre l'acte illocutoire et l'intention du locuteur (Maudet, 2002 : 72). D'ailleurs, les divers travaux adoptant une telle démarche ne rendent compte que des contraintes structurelles liées au dialogue. C'est le cas notamment du modèle structurel de Genève (Moeschler, 1989). En outre, penser que le locuteur « code » un sens qu'il transmet à un auditeur qui le « décode » est une vision simpliste qui fait fi de l'aspect inférentiel de la communication :

« S'il est vrai qu'une partie de la signification d'un message est directement donnée, cela ne représente pas la totalité de la signification » (Maudet, 2002 : 73).

En cela, Grice (1975) offre une vision plus nuancée de la communication que Searle & Vanderveken (1985). D'ailleurs, la sémantique formelle de ces derniers limite l'étude aux énoncés dont l'interprétation est littérale. Et même s'il est possible de reconnaître l'acte illocutoire grâce à la forme linguistique de l'énoncé, y a-t-il pour autant un lien direct entre un acte et son intention sous-jacente ?

4.2. Deux niveaux

Pour sortir de cette impasse, nous proposons donc de distinguer deux niveaux :

- **Le niveau cognitif** : ce niveau concerne les stratégies et les plans² qu'un individu élabore pour atteindre son but. Si les processus cognitifs ne sont pas directement observables, comme le fait remarquer Brassac (2001), nous pensons qu'ils le sont indirectement sur le plan linguistique par des **marques non décidables**.
- **Le niveau de l'énonciation** : ce niveau concerne l'ancrage du raisonnement cognitif et le rapport qu'il entretient avec la situation d'énonciation. Il donne des indications sur la valeur illocutoire des énoncés, grâce aux **marques décidables**.

C'est pour rendre compte de ce niveau cognitif et pour éviter toute confusion entre l'acte illocutoire (niveau de l'énonciation) et l'intention du locuteur (niveau cognitif), que nous adoptons, comme point de départ de notre cadre théorique, le postulat selon lequel il existe des inférences stéréotypiques récurrentes au sein des ALI.

5. Niveau cognitif et marques non décidables

5.1. Des inférences stéréotypiques récurrentes au sein des ALI

La mise en évidence d'inférences stéréotypiques est une étape nécessaire pour caractériser les sèmes de l'arrière-plan représentationnel, et pour répertorier les marques non décidables. C'est pour répondre à ce besoin théorique que nous avons élaboré dans de précédents travaux (Anquetil, à paraître) une typologie de ces inférences. Celle-ci a été mise au point en répertoriant ce que nous appelons les arguments illocutoires récurrents dans les énoncés supports d'ALI. L'idée est que le sens littéral de ces énoncés contient un argument qui permet de tirer une conclusion et le passage de l'un à l'autre n'est possible que par le biais d'une inférence, d'un topos³. C'est le cas de l'énoncé 6 :

6. Peux-tu atteindre le livre qui est sur l'étagère du haut ?

En effet, le sens littéral de cet énoncé est une question adressée à l'interlocuteur sur ses capacités physiques : l'interlocuteur est désigné ici comme le réalisateur potentiel (R) de l'acte x^4 , soit atteindre le livre sur l'étagère. La réponse à cette question étant donnée par le contexte énonciatif, nous pouvons extraire un argument illocutoire (AI) : R a la capacité physique requise pour atteindre le livre. De cet argument, il est possible de déduire la conclusion illocutoire (CI) implicite : il y a des chances que R accomplisse x , soit atteindre le livre. Et le passage de AI à CI n'est possible que par le biais d'un topos : // Plus R a les facultés physiques requises pour la réalisation de x , plus il y a de chances que R réalise x //.

² Les plans sont les premières structures auxquelles les chercheurs (Allen & Cohen, 1980 et Cohen et Perrault, 1979) ont pensé pour mener à bien leur volonté de formaliser les notions intentionnelles et pour rendre compte des processus cognitifs.

³ Les topoï sont des inférences générales et graduelles, communes et partagées par un groupe social, qui autorisent le passage d'un argument A à la conclusion C (Ducrot, 1988).

⁴ Il s'agit de l'acte visé par le locuteur, soit l'acte perlocutoire.

5.2. « Pouvoir » : marqueur de l'arrière-plan représentationnel

Trois grandes classes composent cette typologie des inférences stéréotypiques présentes dans les ALI. Mais par souci de concision, nous n'évoquerons que la première, à savoir les inférences stéréotypiques qui concernent le patrimoine du réalisateur potentiel (R) de l'acte x. Pour résumer, le topos-type qui caractérise cette classe est //Plus le patrimoine de R est riche, plus il a la possibilité d'agir//. Ce patrimoine comprend le capital physique du réalisateur potentiel de l'acte x, mais aussi son capital mental, économique, social et culturel. Pour chaque forme de capital, nous avons saisi les manifestations linguistiques qui témoignent de cet arrière-plan représentationnel dont voici un bref inventaire :

Formes de capital	Verbes et locutions verbales	Exemples d'énoncés supports
Capital physique	<i>Pouvoir, être assez grand pour, être capable de, entendre, voir ...</i>	<i>Pourrais-tu / Serais-tu assez grand pour atteindre le livre qui est sur l'étagère du haut ? (requêtes)</i>
Capital mental	<i>Pouvoir, penser, comprendre, avoir la force psychologique ...</i>	<i>Tu ne peux pas le laisser dehors ! (directif)</i>
Capital économique	<i>Pouvoir, avoir, posséder, bénéficier de ...</i>	<i>Peux-tu payer ? / As-tu de quoi payer ? (requête)</i>
Capital social	<i>Être ou Avoir + relation sociale entretenue avec L, Connaître ...</i>	<i>Aurais-tu un ami doué en mathématiques ? (requête)</i>
Capital culturel	<i>Savoir, connaître, avoir la compétence ...</i>	<i>Peux-tu / Sais-tu répondre à cette question ? (requête)</i>

Tableau 1 - Inventaire de quelques mots-témoins caractéristiques de la classe 1.

Cette étape initiale de notre recherche rend compte d'un premier phénomène. Le verbe *pouvoir*, qui est le mot-témoin le plus représentatif de la première classe, est un marqueur de dérivation illocutoire qui a la faculté de déclencher un topos ou une inférence stéréotypique se résumant à [+pouvoir ; +agir] : il est la manifestation linguistique du raisonnement cognitif du locuteur. Néanmoins, bien que ce marqueur participe au décodage de l'énoncé, il n'en détermine pas pour autant sa valeur illocutoire. D'ailleurs, les énoncés 7, 8, 9 et 10,

- 7. Je ne **peux** pas faire cela.
- 8. Vous ne **pouvez** pas faire cela.
- 9. Comment **pouvez**-vous faire cela.
- 10. Vous **auriez** quand même **pu** faire cela.

véhiculent la même inférence stéréotypique [+pouvoir ; +agir] que l'énoncé 11, mais leurs valeurs illocutoires diffèrent (directifs, expressifs...) :

- 11. **Pouvez**-vous faire cela ?

Ce phénomène remet en cause l'idée préconçue selon laquelle le verbe *pouvoir* est présent uniquement dans les énoncés servant à accomplir des actes directifs. C'est d'ailleurs le problème que pose l'analyse de Caelen & Fréchet (1994), qui considèrent que les scripts contenant le verbe *pouvoir* sont nécessairement des actes directifs. C'est ce premier constat qui nous conduit à dire que le verbe *pouvoir* est une marque non décidable.

6. Ancrage des inférences stéréotypiques dans l'énonciation et marques décidables

Pour confirmer l'hypothèse selon laquelle seule l'actualisation d'une inférence stéréotypique en contexte permet de déterminer la valeur illocutoire d'un énoncé, nous avons eu recours au logiciel C-Oral-Rom et à la base catégorisée de Frantext. C-Oral-Rom est un corpus multilingue (français, italien, portugais, espagnol) de parole spontanée composé d'environ 1200000 mots, soit quelque 121 heures de parole recueillies en milieu naturel. Le but visé par les initiateurs de ce logiciel est de représenter la variété des actes de parole dans la

langue de tous les jours, et de faciliter l'induction des structures prosodiques et syntaxiques d'un point de vue quantitatif et qualitatif. Cette ressource a été conçue pour la modélisation prosodique, les procédures de test en HLT (Human Language Technologies) et l'étude de la parole spontanée basée sur les corpus. Le corpus en langue française de C-Oral-Rom est de 269959 mots. Le second outil utilisé pour notre recherche, Frantext, est quant à lui constitué d'un vaste corpus de textes littéraires français et d'un logiciel offrant une interface Web. Deux versions de Frantext sont proposées. La première, la base non catégorisée, comprend 4000 textes, plus de 210 millions d'occurrences et environ un million d'auteurs. Les textes appartiennent aux domaines des sciences, des arts, de la littérature, des techniques et couvrent cinq siècles de littérature (du seizième au vingtième siècle) : 80 % de ces textes appartiennent à des œuvres littéraires, et 20 % à des œuvres scientifiques. La seconde version, la base catégorisée de Frantext (celle que nous utilisons), recense 1940 textes en prose des XIX^e et XX^e siècles, soit environ 127 millions d'occurrences qui ont fait l'objet d'un codage grammatical selon les parties du discours. Aux fonctionnalités du Frantext intégral ont été ajoutées des possibilités de requêtes portant sur les codes grammaticaux, ce qui présente un avantage considérable pour notre recherche. Le corpus que nous avons pu constituer grâce à Frantext (dans les romans et les œuvres théâtrales) est de 66 027 794 mots et de 945 textes. Grâce à ces outils, nous avons pu tester l'inclinaison de certaines combinaisons de marqueurs à détecter des énoncés supports d'ALI – inclinaison que nous avons observée lors d'études antérieures. Les combinaisons que nous testons constituent quatre structures, dont voici les spécificités linguistiques :

Structure n° (ST)	Adverbe interrogatif	Verbe pouvoir (principale)	Sujet	Mode	Temps	Type de phrases	Ordre des mots	Contenu propositionnel	Type d'actes illocutoires
ST1	Non	Oui	2 ^e pers.	Cond. Ou Ind.	Présent	Interr.	Sujet inversé pour l'écrit ⁵	Non défini	Directifs
ST2	Non	Oui	2 ^e pers.	Cond.	Passé	Décl.	Sujet + verbe	Non défini	Expressifs
ST3	<i>Comment</i>	Oui	2 ^e pers.	Cond. Ou Ind.	Présent ou futur	Interr.	Sujet inversé	Non défini	Expressifs
ST4	Non	Oui	2 ^e pers.	Ind.	Présent ou futur	Décl.	Sujet + verbe	Verbes de certitude ⁶	Promissifs

Tableau 2 - Combinaisons de marqueurs explorées.

Ces structures nous permettent d'obtenir un corpus de 1416 occurrences à partir de Frantext : 775 occurrences pour ST1 ; 278 pour ST2 ; 286 pour ST3 et 77 pour ST4. Le corpus constitué à l'aide de C-Oral-Rom ne présente que 89 occurrences pour la structure 1, dont seulement 43 d'entre elles sont exploitables (les 46 autres ne répondant pas à tous les critères préalablement sélectionnés). Notons aussi que C-Oral-Rom ne nous a pas permis de répertorier un nombre suffisant d'occurrences correspondant aux structures 2, 3 et 4, pour pouvoir envisager un calcul statistique réellement intéressant. Nous attendons de cette recherche qu'elle nous permette de :

⁵ C-Oral-Rom montre qu'à l'oral, il n'y a généralement pas d'inversion du sujet dans les interrogatives. C'est pourquoi, l'ordre des mots de la structure 1 varie selon les corpus d'étude.

⁶ [*Etre certain(e)(s), compter sur ou compter que, être sûr(e)(s), avoir la certitude*] + expression de la première personne ; et *y compter*.

- Démontrer qu'une combinaison de marqueurs constitue de façon quasi-systématique (environ 95 % des cas) des énoncés à même de produire des ALI à valeur illocutoire déterminée.
- Faire émerger d'autres marques linguistiques favorisant la dérivation illocutoire.

Pour atteindre ces objectifs, nous avons créé un concordancier sous Excel et nous avons organisé l'exploitation de notre corpus en trois étapes :

- **Etape 1 :** Classement de chacun des énoncés de notre corpus selon deux axes : les emplois non dérivés, vs. dérivés correspondant à la valeur illocutoire attendue.
- **Etape 2 :** Etiquetage de chacun des énoncés en fonction de marqueurs non sélectionnés préalablement et pouvant décider de la valeur illocutoire des actes illocutoires : présence ou non d'opérateurs argumentatifs, position syntaxique de ces opérateurs, sémantisme du verbe contenu dans la proposition subordonnée.
- **Etape 3 :** Mise en correspondance de la valeur illocutoire de ces énoncés et de ces marques linguistiques récurrentes.

Trois phénomènes linguistiques ressortent de cette étude.

6.1. *Les marques d'énonciation : des marques décidables*

Tout d'abord, les marques décidables sont les marques linguistiques qui ancrent une inférence stéréotypique dans une situation d'énonciation : il s'agit, des pronoms personnels, du temps et du mode du verbe principal, du type de phrases, de l'ordre des mots, et de l'intonation. Sur ce point, nous nous opposons à Vanderveken (1988), pour qui le temps et la personne du verbe ne font que déterminer les contenus propositionnels des énonciations, mais n'en affectent pas les forces illocutoires. Pour nous, ces marques, une fois combinées aux marques non décidables permettent de détecter des énoncés supports de façon automatique, et cela de façon tout à fait satisfaisante pour les ALI les plus conventionnels. Par exemple, avec la structure 1, nous détectons des actes directifs dans 100% des cas à l'oral et dans 92,13% des cas à l'écrit. Si l'on concentre notre étude sur la période allant de 1950 à 2000, le résultat obtenu à l'écrit passe à 96,38 % pour les romans. Observons au passage que le simple fait de remplacer une des marques d'énonciation par une autre, modifie la valeur illocutoire de l'énoncé. En effet, l'ordre des mots (13), la personne (14), le temps (15), la négation (15) et le mode du verbe de la principale (16) ont une incidence sur la valeur illocutoire des énoncés :

12. **Peux-tu** me les avancer, ces dix millions ? (requête)
13. **Tu peux** me les avancer, ces dix millions. (assertion)
14. **Je peux te** les avancer, ces dix millions. (assertion et proposition)
15. Tu **ne pouvais pas** me les avancer, ces dix millions ? (reproche)
16. Tu **aurais pu** me les avancer, ces dix millions ? (reproche)

Seul le contenu propositionnel de l'énoncé ne semble pas faire varier la valeur illocutoire de l'énoncé, comme nous pouvons l'observer en 17 :

17. **Peux-tu m'expliquer comment c'est arrivé ?** (requête)

C'est d'ailleurs le point de vue que défend Vanderveken (1988). Néanmoins, ce phénomène, ne s'applique, à notre avis, qu'aux actes de langage directs, et aux ALI les plus conventionnels (ce qui est le cas de la structure 1). C'est ce que nous démontrerons dans ce qui suit.

6.2. *Contenu propositionnel et dérivation illocutoire*

Considérer que le verbe *pouvoir* est à lui seul un marqueur de dérivation illocutoire n'a pas de sens. C'est pourquoi nous pensons que des précisions doivent être apportées aux travaux de Roulet (1980). En effet, il apparaît que c'est dans un contexte linguistique spécifique que le verbe *pouvoir* dérive le sens littéral de l'énoncé. C'est d'ailleurs pour cela que les critères

de sélection que nous avons choisis pour constituer le corpus de la structure 3⁷, n'aboutissent pas à des résultats aussi satisfaisants que nous le pensions : près de 92% tout de même dans les œuvres théâtrales, mais seulement 86,22% dans les romans. Pour parvenir à détecter de façon systématique des actes expressifs, nous avons donc dû réintroduire de nouveaux critères de sélection. Il s'avère que le sémantisme du verbe contenu dans la proposition infinitive de cette structure participe lui aussi à la dérivation illocutoire. En effet, les verbes renvoyant aux connaissances et aux compétences ou du moins, à un savoir-faire particulier, annulent dans la plupart des cas la dérivation illocutoire que produit le verbe *pouvoir* (dans 90% des cas, en moyenne). Par exemple, l'énoncé 18 a une signification purement littérale :

18. Comment peux-tu nous **prouver** cela ?

Par conséquent, si grâce à un filtre, nous écartons ce type de verbes de notre corpus, nous obtenons 95,74% d'actes expressifs pour les romans, et 94,91% pour le théâtre, soit un écart de près de 10% pour les romans. D'ailleurs, nous avons pressenti ce phénomène en choisissant comme critère de sélection des verbes de certitude (pour la structure 4), ce qui nous permet d'identifier à coup sûr des actes promissifs (à 100%). Ainsi, nous pensons que le contenu propositionnel participe à la dérivation illocutoire, puisqu'il est susceptible de faire varier le sens de l'énoncé. En cela, nous rejoignons le point de vue de Fraser, 1975.

6.3. Les opérateurs argumentatifs

Enfin, certains éléments linguistiques, même s'ils ne favorisent pas particulièrement la dérivation illocutoire, confirment la valeur illocutoire de l'énoncé. C'est le cas notamment d'adverbes ou de locutions adverbiales et prépositionnelles portant sur le verbe de la principale (bien souvent le marqueur de dérivation illocutoire) comme *au lieu de*, *bien*, *au moins*, *quand même*, *tout de même*. En effet, leur simple présence dans l'environnement linguistique du verbe de la structure 2 certifie à 100 % que l'énoncé est un acte expressif. De même, si *mais* est antéposé à la structure 2, l'énoncé aura la valeur d'un acte assertif dans 90,90% des cas, alors que si *mais* est postposé à cette même structure, l'énoncé sera un acte expressif dans 87,5% des cas. Ainsi, la présence de marqueurs dans l'environnement linguistique permet de confirmer la valeur illocutoire des énoncés.

7. Conclusion

La formalisation des ALI constitue un enjeu important pour les interactionnistes et pour les chercheurs en TAL. D'où l'intérêt de cette recherche : elle nous éclaire sur le fonctionnement syntaxique des énoncés supports d'ALI. Le pivot de ces énoncés est généralement un marqueur dit non décidable : il renvoie à l'arrière-plan représentationnel propre aux ALI, et plus précisément à une des inférences stéréotypiques constitutives de cet arrière-plan. Et c'est l'ajustement contextuel de cette inférence qui détermine la valeur illocutoire de l'énoncé : ajustement qui est d'ordre actanciel, temporel, modal, situationnel, etc. Cet ajustement est marqué linguistiquement par des marques dites décidables. Si ces combinaisons de marqueurs permettent de détecter des énoncés supports d'ALI, c'est que ces tournures revêtent une forme de figement – figement qui est l'indice d'une association conventionnelle entre représentation formelle et représentation sémantique. Toutes ces données syntaxiques et sémantiques, correctement formalisées, constitueront le programme informatique composant le système expert, nécessaire aux chercheurs en TAL pour élaborer des outils de traitement automatique (de corpus ou de traduction). Dès lors, diverses applications pratiques seront possibles : analyse d'un phénomène discursif donné (requête, reproche etc.) dans des volumes textuels considérables et variés ; étude des codes

⁷ Notons que cette structure n'est pas fréquente à l'oral. Aucun énoncé de ce type n'a été répertorié dans C-Oral-Rom.

sociolinguistiques propres aux routines conversationnelles en classe de FLE ; illustrations didactiques d'un des problèmes les plus courants en traductologie (la traduction des ALI). Pour poursuivre la réflexion, il serait intéressant :

- D'effectuer une recherche analogue sur des énoncés moins conventionnels. De cette façon, nous pourrions montrer l'influence des données contextuelles sur le sens de l'énoncé et voir si ces données sont captables en termes de marques linguistiques.
- D'établir une hiérarchie de ces marqueurs, en fonction de leur degré d'impact sur la valeur illocutoire des énoncés.
- De calculer le pourcentage d'énoncés supports d'ALI identifiés dans un corpus donné afin de pouvoir évaluer la couverture de ce modèle formel.

Références

Références bibliographiques

- ALLEN, James, PERRAULT, Raymond C. (1980). Analysing intention utterances, *Artificial intelligence*, 15 : 143-178.
- ANQUETIL, Sophie (à paraître). Actes de langage indirects et inferences stéréotypiques : le patrimoine, un argument illocutoire, *Actes du colloque "Stéréotypage, stereotypes : fonctionnements ordinaires et mises en scène. Perspectives interdisciplinaires" des 21, 22 et 23 juin 2006*, Université Paul-Valéry, Montpellier III.
- BRASSAC, Christian (2001). L'interaction communicative, entre subjectivité et interagentivité, *Langages*, 144 : 39-57.
- CAELEN, Jean, FRÉCHET, Anne-Lise (1994). Attitudes cognitives et actes de langage en situation de communication homme-machine, *Cahiers de l'ICP, Rapport de Recherche 3* : 137-152.
- COHEN, Philip R., PERRAULT, RAYMOND C. (1979). Elements of a plan-based theory of speech acts, *Cognitive science*, 3 : 177-212.
- DUCROT, Oswald (1988). Topoi et formes topiques, *Bulletins de linguistique française*, 22 : 1-14.
- FRASER, Bruce (1975). Hedged performatives, in *Syntax and semantics 3*, Cole, Peter & Morgan, Jerry (Eds). New York : Academic Press, 187-210.
- GRICE, Herbert Paul (1975). Logic and Conversation, in *Syntax and semantics 3*, Cole, Peter & Morgan, Jerry (Eds). New York : Academic Press, 64-75.
- MAUDET, Nicolas (2002). À la recherche de la structure intentionnelle dans le dialogue, *Traitement Automatique des langues*, 43, n°2 : 71-98.
- MOESCHLER, Jacques (1989). *Modélisation du dialogue*. Paris : Hermès.
- ROULET, Eddy (1980). Modalité et illocution, *Communications*, 32 : 216-219.
- SEARLE, John R. (1982). *Sens et expression*. Paris : Minuit.
- SEARLE, John R., VANDERVEKEN, Daniel (1985). *Foundations of illocutionary logic*. Cambridge : CUP.
- VANDERVEKEN Daniel (1988). *Les actes de discours, Essai de philosophie du langage et de l'esprit sur la signification des énonciations*. Liège, Bruxelles : Pierre Mardaga.

Références aux sites Internet

ATILF. Base catégorisée de FRANTEXT,
<http://www.atilf.atilf.fr/frantext.html> Consulté en septembre 2006.

Références aux logiciels

CRESTI Emanuela, MONEGLIA Massimo (eds), (2005). *C-Oral-ROM, Integrated Reference Corpora for Spoken Romance Languages*. Studies in Corpus Linguistics 15. Amsterdam Philadelphia : John Benjamins.

CONSTITUTION DE RESSOURCES PÉDAGOGIQUES NUMÉRIQUES : LE LEXIQUE DES AFFECTS

Magdalena AUGUSTYN, Sabrina BEN HAMOU, Gwendoline BLOQUET, Vannina GOOSSENS, Mathieu LOISEAU, Fanny RINCK

magdalena.augustyn@bvra.e.u-grenoble3.fr, sabrina_ben_hamou@yahoo.fr,

bloquetgwen@yahoo.fr, vannina.goossens@u-grenoble3.fr,

mathieu.loiseau@u-grenoble3.fr, fanny.rinck@u-grenoble3.fr

LIDILEM, Université Stendhal Grenoble 3

BP 25 - 38040 Grenoble cedex 9

Abstract : The work presented in this article is part of a project aiming at the development of French teaching pedagogical resources and their NLP¹-based exploitation. The article focuses on the “affect lexicon” to exemplify the process of computer modelling of textual resources (the description of the lexicon and its semantic annotation within the corpus) on which the extraction of examples and the generation of exercises are based.

1. Introduction

Le travail présenté s’inscrit dans un projet² qui vise la constitution et l’exploitation de ressources linguistiques pour la didactique du français, à destination d’un public de français langue étrangère (niveau intermédiaire à avancé) et d’un public de français langue maternelle (collège et lycée). Les ressources, textuelles, sont traitées à l’aide d’outils de traitement automatique des langues (TAL), et servent l’enseignement et l’apprentissage de phénomènes énonciatifs, liés à l’expression du point de vue et de la subjectivité dans les textes.

Nous nous centrerons ici sur le travail préalable à l’exploitation pédagogique, et relatif au lexique des affects : le ressenti ou l’attitude du narrateur et/ou de ses personnages sont caractérisés lexicalement, comme relevant de la « joie », de la « tristesse », etc.

Le projet réunit des linguistes, des spécialistes du TAL et des didacticiens, et prend comme point de départ l’idée qu’une approche en corpus de ces phénomènes doit favoriser leur enseignement/apprentissage. Il s’agit de recourir à des données langagières attestées, qui permettent aux apprenants de repérer, d’analyser et de manipuler ces phénomènes, leurs co-occurrences et leurs effets en termes d’expression de la subjectivité. Une telle approche en corpus doit ainsi favoriser la réflexion métalinguistique des apprenants, dans une perspective en particulier de réinvestissement des apprentissages lexicaux dans la production de textes. Cet objectif requiert d’une part une modélisation informatique des ressources textuelles : il s’agit de faire en sorte qu’elles soient utilisables en tant que ressources pédagogiques, que ce soit comme réservoir d’exemples pour l’enseignant, qui doit pouvoir formuler des requêtes dans le corpus, ou comme réservoir d’exercices pour l’apprenant. D’autre part, une description linguistique du lexique des affects doit être faite de manière à en faciliter l’enseignement et l’apprentissage. Elle doit permettre d’annoter le phénomène dans les textes du corpus, ainsi enrichi d’informations sémantiques qui sont exploitées automatiquement pour la formulation de requêtes dans le corpus et la génération d’exercices.

¹ Natural Language Processing

² Plan pluri-formation piloté par le Lidilem (Université Stendhal Grenoble 3).

Ce sont ces deux aspects, la modélisation informatique des ressources et la description du lexique des affects à des fins d'annotation sémantique que nous présentons ici.

Le corpus utilisé représente 850.000 mots. Il comprend des textes littéraires principalement, en raison du public visé ainsi que des contraintes imposées par la diffusion des textes. L'intérêt qu'il y aurait à disposer d'autres types de textes, notamment des textes journalistiques et des textes en usage sur Internet se heurte à des problèmes de droit. On peut également se demander dans quelle mesure le contenu du web peut servir de corpus, a fortiori à des fins d'exploitation pédagogique.

2. Modélisation informatique des ressources

La mise à disposition pour des enseignants de français de ressources pédagogiques numériques fait intervenir diverses compétences : didactiques (exploitation des ressources constituées), linguistiques (description des phénomènes linguistiques) et informatiques (numérisation et mise à disposition des ressources). Tous les aspects de ce projet sont interdépendants. La description des phénomènes linguistiques nécessite d'avoir des informations d'ordre didactique (comment les enseignants voudront-ils exploiter les informations linguistiques mises à leur disposition, pour quel type d'activité ?) et relevant du TAL³ (dans quelle mesure l'annotation des phénomènes pourra-t-elle être automatisée ?). De leur côté, les didacticiens ne peuvent préparer l'exploitation pédagogique sans une connaissance préalable approfondie des ressources linguistiques disponibles (quelles informations seront disponibles dans les ressources et comment elles seront présentées ?).

D'un point de vue informatique, le processus de constitution du corpus est lui aussi tributaire des choix de modélisation linguistique des phénomènes et sera détaillé dans la partie suivante à travers l'exemple du lexique des affects. De la même manière, la création du système *via* lequel les utilisateurs exploiteront les ressources doit prendre en compte leurs attentes et dépend des informations contenues dans ces ressources.

Toutes ces contraintes nécessitent un développement incrémental pour chaque notion incluse dans les ressources développées. Ce type de fonctionnement requiert une architecture globale du système la plus générique possible. Nous utilisons donc l'architecture de la plateforme d'apprentissage des langues MIRTO⁴, développée à l'université Stendhal. La modularité de la plateforme provient de la dissociation des données (corpus) et des traitements (Antoniadis *et al.*, 2004a) et permet ainsi d'éviter les principaux défauts des logiciels d'apprentissage des langues : la pauvreté des informations de sens associées aux données langagières (Antoniadis *et al.*, 2004b) et le manque de diversité des exemples disponibles.

Dans cette architecture (cf. Figure 1), un script (S_i) est un programme informatique qui crée une activité à partir d'un texte donné. Cette activité peut aussi bien être un exercice lacunaire géré entièrement par l'ordinateur (ex : exercice lacunaire sur les noms d'affect) que la présentation de certaines facettes du texte pour un travail dirigé par l'enseignant en présentiel. On peut, par exemple imaginer un script qui, dans un texte donné, mettrait en évidence toutes les occurrences des noms d'affect de la même manière afin de générer, après un travail d'exploration, une discussion aboutissant à la description des différentes catégories de noms d'affect. Dans la première phase du travail, les apprenants surligneraient d'une manière différente chacun des noms d'affect en fonction des types qu'ils remarquent, avant de confronter, dans une seconde phase, la typologie ainsi créée avec la typologie existante. Pour un tel script, les différentes phases du travail seraient configurables par l'enseignant, qui fournirait ainsi le contexte didactique de l'application du script :

³ Traitement Automatique des Langues

⁴ Multi-apprentissages Interactifs par des Recherches sur des Textes et l'Oral

- choix du (ou des) texte(s) au(x)quel(s) appliquer le script
- choix des phénomènes à mettre en valeur (noms d'affect dans l'exemple ci-dessus)
- avec quelle précision (domaine, sous-domaine ou distinction plus fine encore, cf. partie 3)
- interaction avec l'utilisateur (visualisation, regroupement en catégories)
- et enfin la proposition ou non d'une « correction » (comparaison de la catégorisation produite par l'apprenant avec celle effectuée au préalable dans le corpus).

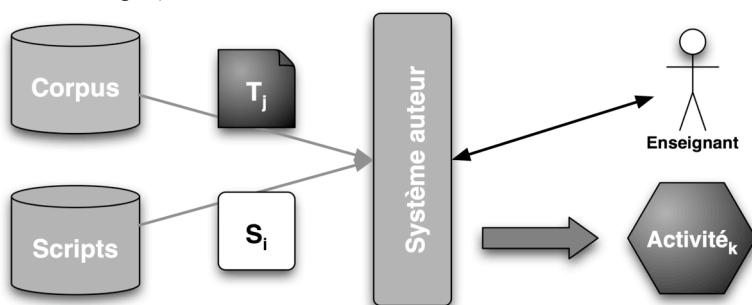


Figure 1 - Génération d'une activité

La création de scripts demandant des compétences pointues en informatique est réservée à un concepteur de scripts, qui proposera, comme nous l'avons vu à travers l'exemple ci-dessus, la possibilité à l'enseignant de définir le contexte didactique de son activité.

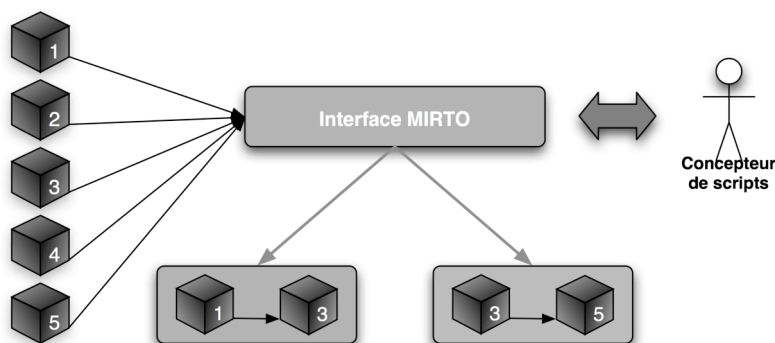


Figure 2 - Conception de scripts dans MIRTO

Via l'interface MIRTO, le concepteur de scripts a accès à un ensemble de traitements de bas niveaux (fonctions numérotées dans la Figure 2). Ces fonctions peuvent relever du TAL (ex : analyse morphologique) ou non (ex : *parser* XML). Le concepteur est alors libre de combiner ces diverses fonctions pour créer des scripts qui seront ensuite accessibles aux enseignants.

Un script de création d'exercices lacunaires portant sur le lexique des affects fera donc intervenir :

- une fonction permettant d'interpréter du XML (*parser* XML),
- une fonction permettant de sélectionner les éléments balisés comme faisant partie du lexique des affects en fonction de paramètres donnés par l'utilisateur enseignant (ex : la catégorie syntaxique des mots, le type, le domaine, cf. partie 3),
- une fonction permettant de créer les trous, qui pourra elle même faire appel à
- une fonction de lemmatisation, selon que l'enseignant (utilisateur du script) voudra ou non afficher les lemmes correspondants aux mots escamotés (en fonction du

niveau des apprenants, l'enseignant choisira entre un exercice lacunaire sans indication ou en affichant sous l'énoncé la liste des formes lemmatisées ou encore la liste des mots tels qu'ils apparaissent dans le texte),

- et enfin une fonction de correction automatique des exercices.

A partir d'un tel script, l'enseignant pourra créer autant d'exercices qu'il y aura de textes dans le corpus, palliant ainsi le manque de diversité des exemples disponibles, récurrent dans les plateformes d'enseignement des langues. Les outils TAL présents sous forme de fonction dans les scripts, ainsi que les annotations préalables du texte enrichissent le sens associé au texte dans le cadre de son utilisation pour l'enseignement du français.

En effet, dans le cadre de ce projet, il ne s'agit pas de travailler avec du texte tout venant mais avec le texte issu du corpus, dont nous allons présenter l'annotation à travers l'exemple du lexique des affects.

3. Le lexique des affects

3.1. Description et organisation du lexique

Le lexique des affects a donné lieu à de nombreuses recherches (voir notamment les numéros 105 de *Langue Française*, Balibar-Mrabti, 1995, et 32 de *Lidil*, Grossmann et Tutin, 2005), sur lesquelles nous nous sommes appuyés pour réaliser des travaux de description linguistique (Tutin *et al.*, 2005, Goossens, 2005) et des travaux à visée didactique (Grossmann et Boch, 2003, Boch *et al.*, 2005, etc.). Tous ces travaux ont mis en évidence d'une part la difficulté qu'il y a à établir quelles unités lexicales font partie du lexique des affects et d'autre part, la complexité sémantique et le caractère polysémique de ces unités lexicales. Pour ce projet, nous avons choisi de retenir une définition assez large de ce qui constitue le lexique des affects afin d'analyser au mieux l'expression de la subjectivité et la polysémie. Nous avons ainsi considéré comme appartenant au lexique des affects les unités lexicales qui peuvent renvoyer à un processus psychologique de type affectif – même s'ils ne renvoient pas obligatoirement à un affect dans tous leurs emplois, comme nous le verrons par la suite. Ceci explique que nous ayons pris en compte des noms comme *bonté* ou *orgueil*, qui s'apparentent en règle générale à des qualités mais qui peuvent également prendre un sens affectif, comme par exemple *bonté* dans le passage suivant, issu de notre corpus :

« Et plus leurs cœurs se rapprochaient, plus ils s'appelaient avec cérémonie "Monsieur et Mademoiselle", plus aussi leurs regards se souriaient, se mêlaient ; et il leur semblait qu'une bonté nouvelle entraînait en eux »
(Maupassant, *Une vie*).

Afin de pouvoir annoter le lexique des affects dans notre corpus nous avons élaboré une liste des unités lexicales que nous considérons comme appartenant au lexique des affects. Nous avons pris comme point de départ des listes existantes, notamment celles de Y. Y. Mathieu (2000) pour les verbes. Dans un souci d'exhaustivité, nous avons alors cherché à élargir ces listes. Nous sommes partis de dictionnaires de référence dans lesquels nous avons navigué par synonymes à partir des unités lexicales de la liste initiale. Nous avons ainsi abouti à l'élaboration de quatre listes distinctes, une pour chaque catégorie grammaticale du lexique des affects : noms, verbes, adjectifs et adverbes. Nous allons nous centrer dans la suite de cet article sur le traitement réservé aux noms.

Afin de décrire et d'organiser notre liste de noms d'affect dans une perspective d'exploitation pédagogique, nous avons élaboré une typologie de ces noms. Nous avons fait le choix d'une typologie notionnelle, ce qui nous a semblé le plus adapté pour l'accès sémantique aux éléments lexicaux par les apprenants. Pour cette même raison, nous avons également fait le choix de nous limiter à deux niveaux hiérarchiques : des domaines, qui sont

au nombre de 27, divisés pour certains en sous-domaines. Il s'agit d'étiquettes conventionnelles regroupant une même « famille » sémantique de noms d'affect. Par exemple, le domaine *joie* se divise en 4 sous-domaines :

Domaine 'joie'			
'Joie'			
'bonheur-satisfaction'	'enthousiasme-entrain'	'joie-gaïeté'	'plaisir-jouissance'
<i>Bonheur, béatitude, contentement, félicité, ivresse...</i>	<i>Alacrité, enthousiasme, optimisme, entrain</i>	<i>Euphorie, liesse, enchantement, joie, gaïeté...</i>	<i>Goût, plaisir, jubilation, volupté, jouissance...</i>

Tableau 1 - Domaine 'joie'

Un certain nombre de noms se trouvent dans plusieurs domaines ou sous-domaines, du fait de la polysémie propre à chaque nom. *Horreur*, par exemple, se trouve à la fois dans le domaine « peur » et dans le domaine « dégoût ». Mais ce n'est pas le seul type de phénomènes polysémiques auquel donnent lieu les noms d'affect et que nous ayons décrit dans ce projet.

3.2. La polysémie des noms d'affect

Comme nous l'avons évoqué précédemment, les noms sélectionnés ne renvoient pas toujours à des affects, ou pas uniquement à des affects. Nous nous trouvons ainsi, en fonction du contexte, en présence de cas de polysémie régulière ou systématique (lorsqu'une dérivation sémantique existe pour tous, ou en tout cas une grande partie, des éléments d'une classe sémantique, en l'occurrence les noms d'affect), ou de polysémie que l'on pourrait appeler irrégulière (lorsque cette dérivation est spécifique à une unité lexicale particulière). Nous avons ainsi distingué quatre cas de figure différents, que nous allons expliquer et exemplifier.

Tout d'abord, et ceci représente évidemment les cas les plus courants, le nom peut exprimer un affect. Par exemple, dans :

« Parfois, si les préjugés ou les instincts de son pays revenaient l'assaillir et lui montraient une vengeance facile au détour d'un sentier, il les écartait avec horreur en pensant à ses camarades de régiment, aux salons de Paris, surtout à miss Nevil » (Mérimée, Colomba)

horreur renvoie à un affect du domaine « dégoût ». Dans :

« Le berger s'était emparé du cheval, mais il avait laissé tomber selle et bride, et paraissait saisi d'horreur » (Mérimée, Colomba)

il appartient dans ce cas au domaine « peur ». C'est également le cas pour des noms comme *orgueil*, *bonté* ou encore *tendresse*⁵, comme dans :

« entre eux flottait déjà cette subtile et vague tendresse qui naît si vite entre deux jeunes gens, lorsque le garçon n'est pas laid et que la jeune fille est jolie » (Maupassant, Une Vie).

Ces noms peuvent également exprimer un affect couplé à une autre dimension sémantique spécifique. Dans de très nombreux cas, par exemple, le nom renvoie à la fois à un affect et à une manifestation de cet affect (ou une attitude exprimant cet affect), comme dans :

« Colomba parlait avec enthousiasme de la beauté de miss Nevil » (Mérimée, Colomba) ou encore « Il y eut parmi toute l'étude, un murmure d'admiration » (Daudet, Le petit chose).

⁵ Qui s'apparentent en règle générale à des qualités.

Parfois, ces noms n'expriment pas un affect mais une autre acception, dérivée par une relation régulière. Par exemple dans :

« *sous sa rude moustache, effroi des chiens et des traîneurs de bancs, se cachait un bon sourire attendri* » (Daudet, *Les contes du lundi*)

effroi ne renvoie pas véritablement à un affect mais plus précisément à ce qui cause cet affect, en l'occurrence la moustache. C'est également le cas pour *peine* et *embarras* dans :

« *je suis une peine et un embarras à ma famille* » (Sand, *La petite Fadette*)

Pour donner un autre exemple de ces acceptions polysémiques régulières, le nom peut exprimer une qualité d'une personne. C'est le cas bien entendu de noms comme *orgueil*, *bonté* ou *tendresse* que nous avons évoqués précédemment et qui s'apparentent habituellement à des qualités, mais également de noms qui désignent généralement des affects comme *passion* dans :

« *il est impossible de parler des Corses sans attaquer ou sans justifier leur passion proverbiale* » (Mérimee, *Colomba*).

Enfin, dans certains cas, ces noms n'expriment ni un affect ni une acception polysémique régulière, comme par exemple *blessure* dans :

« *La blessure à la tempe est grave, n'est-ce pas ?* » (Leroux, *Le mystère de la chambre jaune*)

ou encore *bonheur* dans :

« *Elle cherchait le trèfle à quatre feuilles, qui se trouve bien rarement et qui porte bonheur à ceux qui peuvent mettre la main dessus* » (Sand, *La petite Fadette*).

Ces différents cas de figure ainsi que les domaines et/ou sous-domaines sont déterminés et annotés en contexte pour chaque nom, de même qu'un certain nombre d'autres traits que nous allons maintenant présenter.

3.3. L'annotation du lexique des affects

L'annotation a été réalisée en deux étapes et sous la forme de balises XML à l'aide d'outils de traitement de corpus : Nooj (Silberztein, 2002) dans un premier temps, pour l'annotation automatique sans attributs puis le filtrage (par exemple suppression d'erreurs de catégorisation comme *je doute, sans doute...*) ; puis Intex (Silberztein, 1992) dans un deuxième temps, pour l'insertion automatique des attributs. A ce stade toutes les valeurs possibles sont attribuées à chaque nom (tous les sous-domaines par exemple) qui sera enfin vérifié et désambiguïsé manuellement en fonction du contexte. Au final, nous aboutissons à un balisage de ce type :

« *Jean Peyrol devient rouge de* <LEXIQUE TYPE = "affect+manif_attitude" CAT = "N" DOMAINE = "plaisir" NV_LANGUE = "courant" INTENSITE = "moyen" POLARITE = "positif" FIGE = "" FIGURE_STYLE = "">*plaisir* </LEXIQUE> *et s'empresse autour de son illustre visiteur.* » (Daudet, *Le petit chose*).

Le choix des différents attributs annotés pour le lexique des affects et leur description a été déterminé par la perspective de l'exploitation pédagogique du corpus. Certains sont communs à toutes les catégories syntaxiques comme le domaine, le niveau de langue (courant, littéraire, etc.), l'intensité (haut ou moyen) et la polarité (négatif, positif ou neutre), alors que d'autres sont spécifiques à une catégorie comme le figement pour les noms, le type agentif ou causatif pour les verbes.

En ce qui concerne les noms, nous avons tout d'abord un attribut TYPE, qui correspond au traitement de la polysémie que nous avons abordé précédemment. Quatre configurations sont possibles : affect, affect couplé à une dimension sémantique particulière (affect+manif_attitude, etc.), non affect mais acception polysémique régulière (non affect+cause_objet, non affect+qualité, etc.), ou non affect. L'attribut DOMAINE comprend le domaine du nom ou le sous-domaine lorsqu'il y en a un, tels qu'établis dans notre typologie.

Dans le cas de contraintes syntaxiques propres à certaines constructions (comme par exemple *c'était pitié de, faire peine, de peur de*, etc.), nous annotons cet emploi discursif particulier à l'aide de l'attribut FIGEMENT, qui comporte deux valeurs : figé ou non. Cette description sera par la suite affinée par des marqueurs de niveau de figement (collocations, expressions figées, etc. cf. Tutin et Grossmann, 2003).

Enfin, nous annotons également l'emploi stylistique des noms d'affect par l'attribut FIGURE_STYLE avec la désignation du type de procédé employé (métaphore, métonymie, personnification, etc.).

Les verbes, adjectifs et adverbes d'affect ont été annotés selon la même méthode. De même, le corpus contient des annotations pour d'autres phénomènes (discours rapporté, modalisation, etc.) et ce sont ces informations qui sont ensuite interprétées par les scripts de la plateforme MIRTO.

4. Conclusion

Il reste pour finir à dire quelques mots de l'exploitation didactique des annotations faites dans le corpus. La perspective didactique détermine d'abord comme nous l'avons vu le choix des annotations. Des entrées notionnelles ont été privilégiées, pour décrire le lexique des affects à partir de domaines tels que la joie, la tristesse, etc., et il faut gérer un équilibre entre la finesse de la description et la nécessité que les annotations fassent sens pour les enseignants. Un des enjeux de l'annotation a par ailleurs consisté à mettre l'accent sur le problème de la polysémie : le but n'est pas de proposer une interprétation univoque des phénomènes, mais au contraire de défendre l'idée que la polysémie peut et doit faire l'objet d'un traitement en classe de langue.

Outre leur utilisation *via* les scripts (évoquée dans la partie 2), les annotations sont exploitées dans les requêtes que pourra formuler l'enseignant, pour tirer des exemples du corpus, ou même les apprenants, dans le cadre d'une activité d'exploration du corpus encadrée par l'enseignant pour examiner les phénomènes qui constitueront l'objet du cours. Un script pourra s'appliquer ensuite aux textes choisis. D'autre part ces annotations permettent aux apprenants de bénéficier d'exercices générables automatiquement. Dans les deux cas, il s'agit de privilégier une analyse et une manipulation du phénomène en contexte, et de sensibiliser les apprenants aux associations lexicales privilégiées (par ex. *rouge de colère*, etc.). Les ressources doivent permettre un travail sur les effets de sens induits par le contexte, leur diversité et les éventuelles formes d'ambiguïté et de polysémie. Le rôle de l'enseignant apparaît alors comme central dans le déroulement des séances : les outils proposés ne sont pas une fin en soi, mais peuvent par contre être le moteur de nouvelles activités non automatisables, et en premier lieu les échanges au sein de la classe que l'observation de ces phénomènes linguistiques doit susciter.

Une première expérimentation (Cavalla et Tutin, 2006) a été menée dans un cours de FLE de niveau intermédiaire à avancé ; elle valide l'intérêt d'un travail d'analyse en corpus et de mémorisation par des exercices pour l'acquisition des collocations « verbe+nom » et leur réemploi dans une activité de production de textes.

Références

Références bibliographiques

- ANTONIADIS, Georges, ECHINARD, Sandra, KRAIF, Olivier, LEBARBE, Thomas, LOISEAU, Mathieu & PONTON, Claude (2004a). NLP-based scripting for CALL activities, *Actes du workshop de Coling '04 sur eLearning for Computational Linguistics and Computational Linguistics for eLearning*, Genève, Suisse, août 2004.
- ANTONIADIS, Georges, ECHINARD, Sandra, KRAIF, Olivier, LEBARBE, Thomas, LOISEAU, Mathieu & PONTON, Claude (2004b). CALL: from current problems to NLP solutions, *Actes de EUROCALL*, Vienne, Autriche, 1-4 septembre 2004.
- BALIBAR-MRABTI, Antoinette (Ed) (2005). Grammaire des sentiments, *Langue Française*, 105.
- BOCH, Françoise, CAVALLA, Cristelle & GROSSMANN, Francis (2005). Apprendre à intégrer le lexique des sentiments au cours de l'écriture de textes narratifs, *Colloque international Lexique et production verbale : vers une meilleure intégration des apprentissages lexicaux*, 9-10 mars 2005, Lyon, France : Institut National de Recherche Pédagogique.
- CAVALLA, Cristelle & TUTIN, Agnès (2006). Lexique de l'émotion en classe de FLE : une expérimentation basée sur corpus, *7e Conférence Teaching and Language Corpora : TaLC2006*, 1-4 juillet 2006, Paris 7, France.
- CRUSE, D. Alan (1986). *Lexical Semantics*. Cambridge : University Press.
- GOOSSENS, Vannina (2005). Les noms de sentiment : esquisse de typologie sémantique fondée sur les collocations verbales, *Lidil*, 32:103-121.
- GROSSMANN, Francis & BOCH, Françoise (2003). Production de textes et apprentissage lexical : l'exemple du lexique de l'émotion et des sentiments, *Repères*, 28: 117-135.
- GROSSMANN, Francis & TUTIN, Agnès (Eds) (2005). Sémantique des noms et adjectifs d'émotion, *Lidil*, 32.
- MATHIEU, Yvette Yannick (2000). *Les verbes de sentiment : de l'analyse linguistique au traitement automatique*. Paris : CNRS.
- NUNBERG, Geoffrey & ZAENEN, Annie (1997). La polysémie systématique dans la description lexicale, *Langue Française*, 113: 12-23.
- PUSTEJOVSKY, James (1991). The Generative Lexicon, *Computational Linguistics*, 17, 4: 409-441.
- PUSTEJOVSKY, James (1995). *The Generative Lexicon*. Cambridge : MIT Press.
- TUTIN, Agnès & GROSSMANN, Francis (Eds) (2003). *Les collocations, analyses et traitement*. Amsterdam : De Werelt.
- TUTIN, Agnès, NOVAKOVA, Iva, GROSSMANN, Francis & CAVALLA, Cristelle (2006). Esquisse de typologie des noms d'affect à partir de leurs propriétés combinatoires, *Langue Française*, 150: 32-49.

Références aux logiciels

- SILBERZTEIN, Max (1992). *Intex*, <http://intex.univ-fcomte.fr>, (nd). Consulté en septembre 2006.
- SILBERZTEIN Max (2002). *Nooj*, <http://perso.wanadoo.fr/rosavram/pages/noojpag.html>, (nd). Consulté en septembre 2006.

CONTRIBUTION DES CORPUS À L'ENRICHISSEMENT DES DICTIONNAIRES BILINGUES GÉNÉRAUX. APPLICATION AU DOMAINE DE LA VOLCANOLOGIE

Cécile FRÉROT^a, Amélie JOSSELIN-LERAY^b

Cecile.Frerot@u-grenoble3.fr, josselin@univ-tlse2.fr

^aUniversité Stendhal - Grenoble 3

BP 25 - 38040 Grenoble Cedex 9

^bERSS & Université Toulouse Le Mirail

5 Allées Antonio Machado

31058 Toulouse Cedex 9

Abstract : This paper describes how a popular-science corpus can improve the treatment of terms in a general bilingual dictionary if used during the dictionary-making process. It compares the data taken from two existing bilingual dictionaries to those found in a corpus concerning two verbs belonging to the field of volcanology: the English verb *erupt* and the French verb *se réveiller*.

1. Introduction

L'utilisation de corpus en traduction est une perspective séduisante qui attire un nombre croissant de professionnels de la langue, donnant lieu à des études et des réflexions depuis environ une dizaine d'années. Que les travaux menés à partir de corpus viennent enrichir des réflexions théoriques sur le processus même de la traduction (Baker, 1998), qu'ils constituent le socle de descriptions linguistiques contrastives ou bien qu'ils soient utilisés dans des études empiriques à des fins didactiques (Zanettin, 1998 ; Bowker, 1998), l'intérêt grandissant porté au corpus est indéniable.

Par ailleurs, l'avènement de la linguistique de corpus a rendu possible et accéléré l'exploitation de corpus électroniques à des fins lexicographiques. Historiquement, l'utilisation de corpus électroniques en lexicographie monolingue française a été marquée par la confection du *Trésor de la Langue Française* grâce au corpus *Frantext*. Pour les dictionnaires unilingues anglais, le dictionnaire pour apprenant *Collins Cobuild English Dictionary*, basé sur le corpus *The Birmingham Collection of Texts* (Sinclair, 1987), a été pionnier dans son domaine. Ce n'est que bien plus récemment que les corpus ont été introduits pour la confection des dictionnaires bilingues (dictionnaire Oxford-Hachette de 1994, basé sur le corpus *The Bank of French* (Atkins, 1994)).

Dans le contexte de notre étude, nous observons deux éléments :

(i) premièrement, le recours au corpus améliore la qualité globale des traductions par rapport à des ressources plus traditionnelles telles que les dictionnaires. C'est le constat auquel aboutit Bowker (1998) à l'issue d'une étude expérimentale portant sur l'utilisation du corpus en tant que ressource pour le traducteur. Dans cette étude, des apprentis traducteurs ont réalisé d'une part, une traduction en utilisant des dictionnaires généraux et spécialisés, et d'autre part une traduction en exploitant des corpus électroniques constitués pour l'expérience. Les résultats montrent que le recours au corpus améliore la qualité des traductions, notamment au niveau de la compréhension du domaine, du choix des termes et de l'utilisation d'expressions idiomatiques. Ce type d'étude, déjà ancien, est relativement rare ;

(ii) deuxièmement, pour la traduction des textes spécialisés, les traducteurs ne se cantonnent pas à la consultation des dictionnaires spécialisés et des banques de données terminologiques mais utilisent toujours largement les dictionnaires généraux comme le montrent les résultats d'un récent sondage réalisé par Josselin-Leray (2005) auprès de diverses catégories d'utilisateurs de dictionnaires généraux unilingues et bilingues.

Face à ce double constat, nous nous intéressons à la contribution possible des corpus à l'enrichissement des dictionnaires bilingues généraux lors de l'élaboration de ces derniers. Dans la continuité des travaux de Josselin et Frérot (2004), nous confrontons des données issues de dictionnaires généraux bilingues anglais/français à des données extraites d'un corpus de vulgarisation dans le domaine de la volcanologie.

Nous axons nos recherches sur le terme, que nous entendons comme tout mot appartenant à une langue de spécialité, et plus particulièrement sur la catégorie du verbe. En effet, parent pauvre de la terminologie et bien souvent délaissé des études en langue de spécialité (L'Homme, 1998 ; Kübler et Foucou, 2003), le verbe fait l'objet d'un traitement particulièrement lacunaire dans les dictionnaires, notamment du point de vue de son fonctionnement lexico-syntaxique. L'intérêt que nous portons aux verbes et aux collocations verbales est également motivé par les difficultés posées par la phraséologie en langue de spécialité (Roberts, 1994).

Nous caractérisons dans un premier temps les corpus constitués pour l'étude (2.1.) avant de présenter les outils utilisés pour les exploiter (2.2.) ainsi que les dictionnaires analysés (2.3.). Dans un second temps, nous procédons à l'analyse de deux verbes-clé du domaine de la volcanologie : le verbe anglais *erupt* (3.1.) et le verbe français *se réveiller* (3.2.).

2. Description du matériau pour l'analyse

2.1. Corpus

Dans notre étude, nous avons utilisé deux types de corpus : (i) un corpus comparable anglais-français (c'est-à-dire un ensemble de textes rédigés dans deux langues différentes, qui ne sont pas la traduction les uns des autres, mais qui présentent des caractéristiques communes), et (ii) un corpus traduit anglais-français. La complémentarité de ces deux types de corpus n'est en effet plus à démontrer (Altenberg et Granger, 2002 ; Maniez, 2005).

Etant donné que nous souhaitons analyser le traitement microstructurel de certains termes dans des dictionnaires bilingues généraux, les corpus auxquels nous avons décidé de confronter les données lexicographiques sont des corpus de vulgarisation scientifique. Synthétisons tout d'abord les caractéristiques générales des deux corpus¹. Le corpus comparable et le corpus traduit comportent des textes en langue anglaise (variété britannique, américaine et canadienne) et en langue française (variété française et canadienne) portant sur le domaine de la volcanologie. Le corpus comparable comporte environ 400 000 mots par langue, soit un total d'environ 800 000 mots, tandis que le corpus traduit contient environ 100 000 mots par langue, soit 200 000 mots au total. La période couverte par les deux corpus est une vingtaine d'années (de la fin des années 1970 jusqu'aux années 2000). Les textes inclus sont des textes écrits, entiers, de source fiable et que l'on peut qualifier de manière globale de « textes de vulgarisation », même si l'appellation recouvre en réalité plusieurs types de discours (du discours du quotidien généraliste au discours de semi-vulgarisation, en passant par exemple par le discours pédagogique). Les tableaux qui suivent (tableaux 1 et 2) présentent de manière récapitulative les divers niveaux de subdivision du corpus selon le type de texte (nom des subdivisions et taille en nombre de mots) et mettent côte à côte divers exemples de documents anglais et français comparables ou traduits.

¹ Pour plus de détails sur la constitution des corpus, voir Josselin-Leray (2005), pp. 374-403.

Type de document		Ex. de titres français	Nombre de mots	Ex. de titres anglais	Nombre de mots	
TEXTES	Hors Presse	Manuels académiques	<i>Eléments de Géologie</i>	12 689	<i>Essentials of Geology</i>	10 590
		Ouvrages de vulgarisation	<i>Les Volcans et leurs Secrets</i>	81 151	General-Interest Publications of the USGS on Volcanoes	91 208
		Documents du Web	<i>Le Volcanisme de Djibouti</i>	26 732	<i>How Volcanoes Work</i>	21 603
		Exposition	<i>Des Volcans et des Hommes</i>		Ø	
	Presse	Quotidiens généralistes	<i>Le Monde</i> <i>La Presse</i>	145 369	<i>The Guardian (GB)</i> <i>The New-York Times (US)</i> <i>The National Post (CD)</i>	148 642
		Magazines de vulgarisation	<i>Science et Vie</i> <i>Science et Avenir</i>	58 609	<i>New Scientist (GB)</i> <i>Science News (US)</i>	58 633
		Magazines de semi-vulgarisation	<i>Pour la Science</i>	15 689	<i>Scientific American</i>	16 094
		Magazines de voyage / découverte	<i>Géo</i> <i>Terre Sauvage</i>	56 446	<i>Earth National Geographic</i>	54 152
	GLOSSAIRES		3 227			4 686
	TOTAL		400 486			405 608

Tableau 1 - Composition du corpus comparable

Type	Titre	Nombre de mots	Date	Sens de la traduction	Titre	Nombre de mots	Date
Magazine de semi-vulgarisation	Pour la Science	70 132	1979 - 1999	ANG ⇒ FR	Scientific American	67 762	1979 - 1999
Dépêche scientifique	Volcanisme et Volcans d'Auvergne	12 482	2002	FR ⇒ ANG	Volcanism and Volcanoes of Auvergne	11 804	1999
Ouvrage de vulgarisation	Les Feux de la Terre-Histoires de Volcans	28 092	1991	FR ⇒ ANG	Volcanoes-Fire from the Earth	26 268	1993
Total		110 712				105 706	

Tableau 2 - Composition du corpus traduit

2.2. Outils d'exploitation des corpus

L'exploration de ces gros corpus a rendu indispensable l'utilisation d'outils informatisés. Afin d'exploiter le corpus comparable en nous focalisant particulièrement sur la catégorie du verbe, nous avons utilisé un outil d'extraction terminologique et phraséologique, Syntex, qui est opérationnel sur des corpus français et anglais (Bourigault *et al.*, 2005), et qui permet l'extraction de termes simples et complexes, notamment de syntagmes verbaux.

Pour naviguer à l'intérieur même des corpus dans le but de faire des recherches ponctuelles et précises (possibilité que n'offre pas réellement un logiciel d'extraction terminologique),

nous avons eu recours à un concordancier qui nous a permis de faire des recherches plus poussées dans le corpus comparable ; il s'agit du logiciel Cocoon².

Enfin, l'exploitation du corpus traduit a impliqué le recours à un outil capable d'aligner les textes traduits et possédant également des fonctions de concordancier bilingue. Notre choix s'est porté sur l'outil LogiTerm³. Nous tenons à souligner que ces trois outils se sont révélés complémentaires.

2.3. Dictionnaires

Les dictionnaires bilingues anglais-français que nous avons utilisés pour l'étude sont les suivants : *Dictionnaire Oxford-Hachette* de 1996 sur CD-ROM [OXHA], et *Harrap's Shorter* de 2000 sur CD-ROM [HAR]. Ces dictionnaires sont de nomenclature comparable. Le premier a été choisi parce qu'il a été élaboré à partir d'un corpus, le second car il a pour réputation de particulièrement bien traiter les mots scientifiques et techniques.

3. Analyse des exemples

3.1. Le verbe erupt

3.1.1. Choix de l'exemple

Le verbe *erupt* a été désigné comme objet d'analyse pour plusieurs raisons : outre le fait qu'il s'agit d'un terme central du domaine de la volcanologie, il s'agit du verbe le plus fréquent extrait par Syntex (554 occurrences) ; enfin, sa syntaxe est problématique dans une optique traductionnelle. En effet, alors qu'il s'agit d'un verbe simple en anglais, l'équivalent le plus courant en français, *entrer en éruption*, présente un schéma syntaxique plus complexe : V + Prép + N. Ce schéma syntaxique figure bien dans les dictionnaires bilingues analysés.

Toutefois, le parcours des données de la partie anglaise du corpus comparable nous a révélé un fonctionnement du verbe bien plus complexe. Nous avons ainsi rencontré les contextes suivants :

(1A) *In July 1986, just as it appeared that Puu Oo would start the 48th episode, magma began to erupt from new fissures at the base of the cone.*

(1B) *During the subaerial shield building stage aa and pahoehoe erupt from the summit area and rift zones.*

(2A) *After periods of volcanic quiescence lasting as long as hundreds of thousands to even a million years, some Hawaiian volcanoes erupt alkalic volcanic products.*

(2B) *Pillow lavas are volumetrically the most abundant type because they are erupted at mid ocean ridges.*

Les contextes (1A) et (1B) montrent que des noms autres que *volcano*, à savoir *magma* ou encore des types de lave comme *aa and pahoehoe* peuvent être utilisés comme sujets du verbe *erupt*. Or, il n'y a trace de ces noms dans aucun des deux dictionnaires bilingues. De plus, dans les contextes (2A) et (2B), le verbe *erupt* est utilisé de manière transitive, emploi qui n'est mentionné dans aucun des deux dictionnaires bilingues analysés. Enfin, dans ces quatre cas, l'équivalent français du verbe *erupt* n'est pas *entrer en éruption*. Étant donné que les quatre extraits proviennent de la même source, il peut s'agir d'un phénomène idiolectal. Nous avons donc analysé les 554 occurrences du verbe dans la partie anglaise du corpus comparable. Pour établir la proportion de chaque emploi du verbe, nous avons cherché des équivalents potentiels dans la partie française du corpus comparable, puis nous avons analysé

² Concordancier utilisé par le projet du Dictionnaire Canadien Bilingue, à l'Université d'Ottawa (Canada).

³ Développé à Ottawa (Canada) par la société Terminotix.

les 44 occurrences pertinentes de ce verbe dans le corpus traduit, afin d'étudier les traductions possibles. Les résultats chiffrés concernant la répartition de l'emploi transitif / intransitif se trouvent dans le tableau qui suit :

	Emploi intransitif	Emploi transitif
Corpus comparable	512 / 554 (92,4 %)	42 / 554 (7,6 %)
Corpus traduit	43 / 44 (97,7 %)	1 / 44 (2,3 %)

Tableau 3 - Fréquence dans les deux corpus des formes transitives et intransitives du verbe *erupt*

3.1.2. Analyse de l'emploi intransitif

Le tableau 3 révèle que l'emploi intransitif est, de loin, celui qui est le plus répandu.

a. Corpus comparable

Parmi les 512 occurrences du verbe intransitif, 5 ne sont pas pertinentes car elles présentent le cas où le verbe *erupt* est utilisé dans un sens figuré. Nous avons également trouvé 15 occurrences de la forme *erupting volcano*, qui devrait figurer dans les dictionnaires bilingues du fait que la traduction en français présente un schéma syntaxique différent du schéma syntaxique anglais : la forme française est en effet du type N + Prép. + N (*volcan en éruption*).

b. Corpus traduit

Le dépouillement du corpus traduit livre 43 occurrences de la forme intransitive du verbe *erupt* (sous ses diverses formes : *erupt*, *erupts*, *erupted*, *erupting*), ce qui correspond à la quasi-totalité des occurrences.

c. Analyse des noms en position de sujet

L'analyse de tous les contextes où apparaissait le verbe *erupt* nous a permis de rassembler les données qui suivent concernant les sujets du verbe *erupt* intransitif (tableau 4) :

Sujets de <i>erupt</i>	Fréquence dans le corpus comparable	Fréquence dans le corpus traduit
<i>Volcano(es)</i>	176	11
Nom de volcan (<i>Mount St. Helens</i> , <i>Vesuvius</i> , <i>Mauna Kea</i> ...)	166	8
	Sous-total = 342 ⇔ 67 % du nombre total d'occurrences	Sous-total = 19 ⇔ 43,2 % du nombre total d'occurrences
<i>Magma</i>	24	10
<i>Lava</i>	16	6
Divers (<i>molten rock</i> , <i>ash and rock</i> , <i>pyroclastic flows</i> , <i>flood basalts</i> ...)	125	8

Tableau 4 - Récapitulatif des sujets de *erupt* intransitif trouvés en corpus

Les sous-totaux trouvés à la troisième ligne du tableau, ainsi que le fait que la plupart des contextes proviennent de quotidiens généralistes, pourraient expliquer que l'on ne trouve que *volcano* comme sujet du verbe *erupt* dans les dictionnaires généraux analysés. La traduction *entrer en éruption* est dans ces cas-là bien vérifiée en corpus. Toutefois, une analyse plus poussée du corpus traduit nous montre que très nombreux sont les cas où le traducteur effectue une transposition (utilisation d'un substantif à la place d'un verbe) pour traduire *erupt*, comme dans l'exemple suivant :

In 1906 he was actually at Vesuvius when it erupted Il est en 1906 au chevet du Vésuve **en éruption**

Cette possibilité n'est pas évoquée dans les dictionnaires. De plus, les trois dernières lignes du tableau 4 ne devraient pas être mises de côté, car elles représentent une part non négligeable des contextes répertoriés (environ 40%) où, en outre, *erupt* ne peut pas être traduit par *entrer en éruption*. En effet, *magma*, *lava*, et la catégorie que nous avons appelée

« divers » représentent, de manière générale, tout ce qui peut être « éjecté du volcan »⁴. Une recherche d'équivalents potentiels dans la partie française du corpus comparable effectuée en cherchant des contextes associant *magma* ou *lave* et *fissure* a produit notamment les résultats suivants : *Du magma s'épanche de cette fissure ; de la lave rouge-orange sort d'une fissure ; les coulées proviennent d'une fissure*. Le corpus traduit a confirmé que les trois verbes *s'épancher*, *sortir*, *provenir* sont des équivalents possibles de *erupt* intransitif associé aux sujets *magma*, *lave* ou *coulée* et a également livré, entre autres, *s'échapper* et *jaillir*.

3.1.3. Analyse de l'emploi transitif

Rappelons, comme l'indique le tableau 3, que le verbe *erupt* est utilisé de manière transitive dans environ 8% du nombre total d'occurrences du verbe dans le corpus comparable. Ces occurrences proviennent de diverses sources, et correspondent à trois constructions différentes :

1. X erupts Y

Ex : *Volcanoes on other planets, and their moons, do not necessarily erupt the same materials as Earth's volcanoes*

Cette construction est l'usage transitif typique du verbe. Nous en avons trouvé 11 occurrences dans le corpus comparable, et une occurrence dans le corpus traduit. L'analyse du corpus comparable nous a permis d'identifier comme équivalent possible d'*erupt* le verbe *émètre*, comme dans cet exemple : *La fissure émet des laves*. Le corpus traduit a livré *cracher*.

2. Y is erupted (from X / by X)

Ex : *Volcanic material can also be erupted from the side of a volcano to produce a flank eruption*

La deuxième construction correspond à la voix passive. Nous en avons trouvé 15 occurrences dans la partie anglaise du corpus comparable. La partie française nous a livré comme équivalent possible *être émis* comme dans le contexte suivant : *Au total, 2 km³ de lave ont été émis*.

3. Y erupted (from X / by X)

Ex : *Total volume of tephra erupted during the formation of Crater Lake was 18 cubic miles*

Le troisième schéma syntaxique est le plus elliptique, utilisant seulement le participe passé. Nous avons trouvé 18 occurrences de ce schéma dans la partie anglaise du corpus comparable. La partie française nous a livré les contextes suivants : *lave issue de la fissure*, *lave vomie par la fissure*, *laves très fluides se déversant par des fissures de l'écorce*.

3.1.4. Proposition d'entrée de dictionnaire pour le verbe *erupt*

Une ébauche d'entrée de dictionnaire bilingue pour le verbe *erupt* enrichie grâce aux corpus est présentée dans le tableau 5. Cette entrée sépare l'usage transitif de l'usage intransitif⁵. Elle montre comment les équivalents varient selon les sujets et les objets typiques du verbe, et illustre l'utilisation des équivalents fournis, en proposant parfois d'autres procédés de traduction (comme la transposition) dans les exemples, lesquels sont introduits par l'astérisque simple. Les collocations verbales sont introduites par **.

<p>1) V.T. (<i>volcano</i>) entrer en éruption * he observed Vesuvius as it erupted il observa le Vésuve en éruption *Hualalai last erupted in 1801 la dernière éruption du Hualalai remonte à 1801 ** begin to erupt, erupt again se réveiller; (<i>magma, lava</i>) jaillir, s'épancher, sortir * magma began to erupt from new fissures at the base of the cone du magma a commencé à s'écouler depuis la base du cône ; (<i>any other material</i>) s'échapper, jaillir * over a four-year period, hundreds of nuée ardentes erupted from Mt. Unzen's summit area en l'espace de quatre ans, des centaines de nuées ardentes se sont propagées depuis le sommet du Mont Unzen 2) V.I. (<i>magma, lava</i>) vomir, émettre * the speed of a lava flow is determined by the volume of lava that is</p>

⁴ Cet usage est répertorié dans le dictionnaire unilingue anglais *New Oxford Dictionary of English* (2000) qui donne pour un des sens de *erupt* : « be ejected from an active volcano ».

⁵ Pour l'usage transitif, elle veille à présenter les trois constructions possibles.

erupted la vitesse d'une coulée de lave est déterminée par le volume de lave émise ; (*any other material, particularly ash*) éjecter, émettre, cracher * **Mount St. Helens explosively erupted ash** le Mont St Helens a craché de la cendre en explosant ***the volcano is built up by the accumulation of material erupted through the conduit** le volcan se construit par accumulation des débris éjectés par la cheminée * **Volcanoes on other planets do not necessarily erupt the same materials as Earth's volcanoes** les éjecta des volcans d'autres planètes ne sont pas forcément les mêmes que ceux de la Terre.

Tableau 5 - Proposition d'entrée de dictionnaire bilingue pour *erupt*, v.

Passons à présent à l'analyse de notre second exemple : le verbe *se réveiller*.

3.2. *Le verbe se réveiller*

Les termes métaphoriques, comme le verbe *se réveiller* (78 occurrences dans le corpus comparable), sont assez problématiques à traduire car il est difficile pour un utilisateur de savoir si la métaphore se rend à l'identique dans l'autre langue. En ce qui concerne la collocation *volcan / se réveiller*, il n'est pas évident de déterminer quel équivalent est à choisir. Le OXHA propose à juste titre *to become active again*, mais le HAR ne propose aucun équivalent sous *se réveiller*. On trouve cependant *the renewed rumblings ou the awakening of a volcano* sous le terme *réveil* de ce même dictionnaire. La deuxième traduction peut paraître légèrement douteuse étant donné qu'elle ressemble fort à un calque, et il pourrait s'agir d'un phénomène de « translationese ». Une analyse du corpus traduit révèle que le terme *awakening*, ainsi que les verbes *awake/re-awaken*, sont utilisés pour traduire *se réveiller* dans les traductions où le français est la langue-source, comme dans les trois exemples suivants :

[FRA] Dès leur découverte, la question a été posée de savoir si les volcans d'Anvergne pouvaient se réveiller.	[ANG] Since their discovery it has commonly been asked whether the volcanoes of the Auvergne could re-awaken.
[FRA] La vengeance des dieux est terrible, le volcan se réveille. 1184 personnes périrent dans les nuées ardentes et les coulées de boue.	[ANG] The vengeance of the gods was terrible, and the volcano - which had been dormant for six centuries - awoke: 1184 people died in the ash and mud.
[FRA] Hekla est l'un des volcans les plus actifs d'Islande. Depuis son paroxysme de 1104, qui fut, avec celui de l'Oraefajökull en 1362, la plus forte éruption explosive des temps historiques dans ce pays, il a connu 167 réveils, le dernier en janvier 1991.	[ANG] Hekla (above, in a 16th-century print) is one of Iceland's most active volcanoes. Since its eruption in 1104 it has awakened 167 times, most recently in January 1991.

Toutefois, le verbe *reawaken* ou le gérondif *reawakening* sont également utilisés dans les textes anglais qui servent de textes sources, comme dans les exemples suivants :

[ANG] The rapid scientific response to the reawakening of Mount St. Helens in 1980 drew on earthquake and ground movement-monitoring techniques developed at Kilaeua.	[FRA] Les techniques de surveillance des séismes et des mouvements du sol mises au point au Kilaeua ont permis aux géologues de réagir rapidement face au réveil du mont Saint-Helens en 1980.
[ANG] Thus any small-scale activity that might have preceded its reawakening went unnoticed.	[FRA] En conséquence, toute marque d'activité même faible, qui avait peut-être précédé son réveil, passa inaperçue.

L'analyse des données de la partie anglaise du corpus comparable confirme que *reawaken* est utilisé par des locuteurs natifs. De plus, elle permet d'enrichir la liste d'équivalents possibles avec, par exemple : *to be again in eruption, to be about to erupt, to be about to blow, to rumble back to life, to threaten to erupt, to begin to erupt, to erupt again, to erupt into life*.

4. *Conclusions et perspectives*

Dans cet article, nous avons montré à quel point les données extraites de corpus complémentaires (corpus comparable et corpus traduit) peuvent améliorer le contenu des dictionnaires, à la fois en termes de qualité (précision et diversité des équivalents) et en termes de quantité (quantité des termes en langue cible), ce qui peut s'avérer très bénéfique pour le traducteur ou l'apprenti traducteur. En aval de l'objectif premier du présent article,

nous souhaitons à terme intégrer les corpus dans l'enseignement de la traduction spécialisée (i) en exploitant directement les résultats de cette étude, et (ii) en impliquant activement les apprentis traducteurs dans la confrontation des données présentes dans les dictionnaires à celles issues des corpus. L'objectif, très pragmatique, vise à permettre aux apprentis traducteurs d'appréhender au mieux et de « manipuler » de manière très concrète les données des dictionnaires et des corpus afin d'en évaluer l'utilité respective pour la traduction.

Références

- ALTENBERG, Bengt & GRANGER, Sylviane (2002). Recent trends in cross-linguistic lexical studies, in *Lexis in contrast. corpus-based approaches*, Altenberg, Bengt & Granger, Sylviane (Eds). Amsterdam/Philadelphia: John Benjamins, 3-48.
- ATKINS, Beryl T. Sue (1994). A corpus-based dictionary, in *The Oxford-Hachette French dictionary*, Corréard, Marie-Hélène & Grundy, Valérie (Eds). Oxford, Paris : Oxford University Press, Hachette, XIX-XXVI.
- BAKER, Mona (1998). Réexplorer la langue de traduction : une approche par corpus, *Meta*, 43, 4 : 480-485.
- BOURIGAULT, Didier, FABRE Cécile, FRÉROT Cécile, JACQUES Marie-Paule & OZDOWSKA Sylwia (2005). Syntex, analyseur syntaxique de corpus, *Actes des 12èmes journées sur le Traitement Automatique des Langues Naturelles*, Dourdan, France, 17-20.
- BOWKER, Lynne (1998). Using specialized monolingual native-language corpora as a translation resource: a pilot study, *Meta*, 43, 4 : 631-651.
- JSSELIN-LERAY, Amélie (2005). Place et rôle des terminologies dans les dictionnaires généraux unilingues et bilingues. Etude d'un domaine de spécialité : volcanologie, thèse de doctorat, Université Lyon II.
- JSSELIN, Amélie & FRÉROT, Cécile (2004). Corpus-based terminology extraction applied to lexicography: How can a popular science corpus help improve general bilingual dictionaries? in *Practical Applications in Language and Computers, PALC 2003*, Lewandowska-Tomaszczyk, Barbara (Eds). Lodz : Peter Lang, 66-79.
- KÜBLER, Natalie & FOUCOU, Pierre-Yves (2003). Teaching english verbs with bilingual corpora: examples in the computer science area, in *Contrastive linguistics and translation studies*, Granger, Sylviane & Petch-Tyson, Stéphanie (Eds). Amsterdam : Rodopi, 185-206.
- L'HOMME, Marie-Claude (1998). Le Statut du verbe en langue de spécialité et sa description lexicographique, *Cahiers de Lexicologie*, 73, 2 : 61-84.
- MANIEZ, François (2005). La Phraséologie du diagnostic : étude d'un corpus bilingue anglais-français d'articles de recherche dans le domaine de la cardiologie, in *Les Mots de la santé. Influence des sociétés et des cultures sur la formation des mots de la santé*, Beltrán-Vidal, Danièle & Maniez, François (Eds). Lyon : Presses Universitaires de Lyon, Travaux du CRTT : 189-200.
- ROBERTS, Roda P. (1994). Identifying the phraseology of Languages for Special Purposes (LSPs), in *Terminologie et linguistique de spécialité : études de vocabulaires et textes spécialisés*, Edwards, H. Peter & Lapierre, Lise (Eds), Actes de Langue Française et de Linguistique 1994-1995. Halifax : Universitas Dalhousiana, 61-74.
- SINCLAIR, John (1987). *Looking-up – an account of the Cobuild project in lexical computing*. Londres : HarperCollins.
- ZANETTIN, Federico (1998). Bilingual comparable corpora and the training of translators, *Meta*, 43, 4 : 613-630.
- Dictionnaires :
Dictionnaire Hachette-Oxford (1996). Hachette Livre, Oxford University Press (Cd-Rom).
Harrap's Shorter. Dictionnaire anglais-français (2000). Chambers, Harrap (Cd-Rom).

ORGANISATION DES FONCTIONS LEXICALES POUR UN MEILLEUR ACCÈS À L'INFORMATION DANS LE DiCo

Anne-Laure JOUSSE

anne.laure.jousse@umontreal.ca

OLST- Université de Montréal, C.P. 6128, succ. Centre-ville

Montréal (Québec) H3C 3J7

Canada

Lattice-ENS, 1 rue Maurice Arnoux

F-92120 Montrouge

Abstract : This article describes how lexical functions are exploited in order to create a system for the organization of lexical-semantic links within lexical units in the DiCo. Lexical functions seem relevant because they make lexical-semantic links explicit, and because they provide several kinds of linguistic information. Such an organization of lexical functions is shown to be helpful in creating search and browse tools for the DiCo.

1. Introduction

L'apparition du support électronique transforme notre façon de concevoir et d'utiliser les dictionnaires. Ceci affecte autant la manière de décrire les unités lexicales que de naviguer au sein des bases de données lexicales (Pruvost, 1995, 2000 ; Jacquet-Pfau, 2002 ; De Schryver, 2003 ; Zock & Carroll, 2003). Le défi lexicographique consiste donc à dompter le support électronique pour le mettre au profit d'une amélioration de la structuration et de l'accès à l'information. Les bases de données lexicales sont majoritairement conçues selon une structure éclatée qui permet une grande liberté dans la connexion des entités. Comme l'ont montré Steinlin *et al.* (2004) pour le DiCo, une fois totalement atomisée, une base de données peut être accessible par n'importe quel élément et modulable pour multiplier les types de requêtes à l'infini. Après l'atomisation des données, l'étape suivante consiste en leur organisation afin d'offrir des moyens de consultation adaptés. Le mode de présentation des relations lexicales dans une base de données est très étroitement lié à la théorie qui lui est sous-jacente. Parmi les projets de grande envergure, on peut noter WordNet (Fellbaum, 1998) qui bénéficie en arrière-plan d'un certain nombre d'études psycholinguistiques, FrameNet qui repose sur la sémantique des cadres (Fillmore *et al.*, 2003), le DAFLES (Selva *et al.*, 2002), inspiré de la Lexicologie Explicative et Combinatoire ainsi que le DiCo, sur lequel nous travaillons.

Le DiCo, développé à l'OLST (Université de Montréal) par Polguère et Mel'čuk, repose sur les principes de la Lexicologie Explicative et Combinatoire (Mel'čuk *et al.*, 1995), et se concentre sur la mise en valeur et l'explicitation des relations sémantiques (collocations et dérivés sémantiques) entre les unités lexicales. Cette base de données lexicale du français fournit, pour chaque unité lexicale décrite (environ 1000), une courte caractérisation sémantique ainsi que l'ensemble de sa combinatoire lexicale. À la différence de ses modèles concurrents, l'éventail de relations lexicales couvertes est conséquent, tant sur le plan syntagmatique (collocations) que paradigmatisque (dérivés sémantiques). Cependant, ces relations ne sont pas organisées suffisamment pour une consultation interactive du DiCo. Par exemple, sous l'article *couteau*, on aimerait accéder directement à tous les verbes dénotant une action impliquant un couteau (*aiguiser, émousser, manier*, etc.), ou à tous les types de couteau (*à beurre, économe*, etc.). Notre objectif est donc de proposer une organisation des

relations lexicales à partir de laquelle on puisse dériver des modes de consultation de la base de données et permettre un accès plus cohérent à l'information.

Pour ce faire, nous utilisons les fonctions lexicales, qui servent à étiqueter les relations lexicales dans le DiCo. Elles permettent une systématisation dans la description lexicologique dans le sens où elles recensent, explicitent et formalisent les relations sémantico-lexicales entre les collocations et les dérivés sémantiques, *grosso modo* entre les unités lexicales constituant un champ sémantique. Elles se présentent sous la forme d'une formule du type $x \text{ f } y$, où x est le mot-clé et y la valeur (c'est-à-dire, le collocatif ou dérivé sémantique). Par exemple, la fonction **Magn** exprime l'intensité (*rire Magn aux éclats*), **S_{loc}**, l'endroit typique (*abeille S_{loc} ruche*), **CausFunc₁**, la causation d'existence (*soupçon CausFunc₁ éveiller*), etc. Travailler directement sur l'encodage nous permettra de traiter globalement les relations qu'entretient toute unité lexicale avec d'autres. La question est alors de savoir comment, à partir d'un encodage formel des relations lexicales, on peut réussir à dégager une classification multidimensionnelle des relations qui permette de les organiser intuitivement pour la consultation. Nous présenterons dans un premier temps les fonctions lexicales en tant qu'outil potentiel pour la représentation et la structuration des relations lexicales, puis nous détaillerons la façon dont on peut exploiter ce formalisme pour atteindre notre objectif. Pour conclure, nous expliquerons comment nous utiliserons, dans la suite de nos travaux, cette structuration des fonctions lexicales pour créer des outils de consultation du DiCo.

2. Utilisation des fonctions lexicales comme grille de structuration des relations lexicales

2.1. Présentation des fonctions lexicales

Une soixantaine de fonctions lexicales standard¹ simples sont communément admises dans le cadre de la Lexicologie Explicative et Combinatoire. Celles-ci peuvent se combiner entre elles pour représenter des relations plus complexes, par exemple :

/*Causer que ~ existe*/
*assemblée#2 CausFunc₀ convoquer [ART ~]*²
 /* [X] préparer ~pour l'utiliser*/
rasoir PreparReal₁ aiguiser, affiler, affûter [ART ~]

Les fonctions lexicales doivent remplir certaines conditions pour être considérées standard, elles doivent représenter des relations suffisamment larges et abstraites pour pouvoir s'appliquer à un grand nombre d'unités lexicales et donner comme valeur un nombre suffisamment varié d'unités lexicales. Les fonctions lexicales formalisent plusieurs aspects d'une relation lexico-sémantique entre deux unités lexicales : le sens de la relation (**sens**), la partie du discours de la valeur (**pos**), le régime syntaxique de la valeur (**reg.**), la nature paradigmatic ou syntagmatic de la relation (**rel.**). À partir de cet encodage, il est donc possible de dégager des sortes de fiches d'identité des relations lexicales, comme l'illustrent les quatre exemples de la figure 1.

¹ À chaque fonction lexicale est associée une formule de vulgarisation appelée « glose ». On utilisera les conventions d'écritures suivantes : **fonctions lexicales** et les /*gloses*/. Par exemple :

/*Utiliser ~*/ voiture **Real₁**, rouler

² « ART » désigne tout type de déterminant, « ~ » correspond au mot-clé. On lira donc la relation : « *assemblée#2 CausFunc₀ convoquer [ART ~]* » de la façon suivante : « *convoquer* (l'/une/cette) assemblée ».

S₁

ex : *vendre*#1 **S₁** *vendeur*

sens : ‘premier actant du mot-clé’

pos : nom

rel. : paradigmaticque

reg. : --

IncepReal₁

ex : *gare*#a **IncepReal₁** *entrer* [en ~]

sens : ‘commencer à réaliser l'action typique liée au mot-clé’

pos : verbe

rel. : syntagmaticque

reg. : -sujet = premier actant du mot-clé (ici, X = I = *train*)

-objet le mot-clé lui même (ici, *gare*).

S_{loc}

ex : *mouton*#I.a **S_{loc}** *bergerie, enclos*

sens : ‘endroit typique où se trouve ~’

pos : nom

rel. : paradigmaticque

reg. : --

V₀

ex : *haine* **V₀** *hair*

sens : ‘équivalent verbal du mot-clé’

pos : verbe

rel. : paradigmaticque

reg. : possède les mêmes actants sémantico-syntaxiques (X et Y) que le mot-clé

Figure 1 - Illustration de l'extraction d'information à partir des fonctions lexicales

2.2. Enjeu d'une classification des fonctions lexicales

Comme nous l'avons mentionné dans l'introduction, les fonctions lexicales dans le DiCo et dans le DiCouèbe, sa version en ligne (Jousse & Polguère, 2005) ne sont pas organisées intuitivement pour la consultation. Elles obéissent à un classement hétérogène en fonction des différentes parties du discours et du type standard ou non des fonctions. Nous pensons que ce classement est discutable dans la mesure où il ne permet pas une organisation pertinente des données. Plusieurs classifications des fonctions lexicales ont été entreprises, on note notamment celles proposées par Zholkovski et Mel'čuk (1970), Steele et Meyer (1990), Grimes (1990), Alonso Ramos (1993) et alonso Ramos et Tutin (1996). La plupart de ces classifications regroupent les fonctions lexicales selon deux grandes catégories : les fonctions paradigmaticques et les fonctions syntagmaticques. Les premières représentent des relations d'ordre paradigmaticque qu'on appelle les dérivés sémantiques (par exemple, pour *abeille* : *ruche*, *essaim*, *apiculteur*), les secondes, des relations d'ordre syntagmaticque ou collocations (par exemple, pour *abeille* : *piquer*, *bourdonner*).

Nous pensons que les différentes classifications proposées jusqu'ici sont contestables car elles mélangent sur le même plan plusieurs critères de classification : parties du discours (par exemple, les prépositions et verbes), fonctions syntaxiques (par exemple, la modification), forme de conjugaison (par exemple, les formules impératives), sens (par exemple, les verbes supports, les verbes de réalisation) et ne sont confinées que dans un seul axe de classement. Nous attribuons cette hétérogénéité à un manque de perspective d'application. En effet, une organisation des fonctions lexicales n'a de sens que si elle vise une utilisation particulière. Pour mettre à profit les différents types d'information fournie par les fonctions lexicales sans les entremêler, il faut en faire non pas une classification unidimensionnelle mais une classification multidimensionnelle. Nous souhaitons créer une classification des fonctions lexicales qui respecte davantage la façon dont elles sont utilisées dans le DiCo. Cette classification se révélera un meilleur reflet de la façon dont s'organisent les unités lexicales dans le lexique d'une langue.

3. Organisation des relations lexicales selon différentes perspectives

Afin de déterminer les dimensions de classement des fonctions lexicales, nous avons analysé les différents types d'information qu'elles modélisent. Chaque type d'information fera l'objet d'une dimension de classification. Nous en présentons quatre : sens de la relation, partie du discours de la valeur, comportement syntaxique du mot-clé et de la valeur, actants concernés par la relation. Pour élaborer une telle organisation, nous attribuons aux fonctions lexicales un certain nombre de paramètres descriptifs de plusieurs ordres. Nous avons compilé l'ensemble des paramètres associés aux fonctions lexicales dans une grande base de données qu'il sera possible d'interroger, par la suite, au moyen de requêtes multicritères.

3.1. Perspective « sémantique »

À l'instar des dictionnaires thématiques, nous souhaitons proposer une organisation onomasiologique des relations entre les unités lexicales. En d'autres termes, il s'agit de les organiser selon une perspective sémantique pour permettre l'accès à une unité lexicale à partir de son sens. Par exemple, dans l'article *cheval*#I, on pourra atteindre l'unité lexicale *écurie* à partir d'une catégorie générale de type « endroit ». Cependant nous ne voulons pas reproduire les travers d'une classification linéaire figée, même si elle est thématique, mais rester dans une optique multidimensionnelle. L'objectif est de permettre l'accès à une unité lexicale selon ses composantes de sens les plus saillantes. Par exemple, dans la relation suivante :

/~ dure trop longtemps~/

conversation#a **AntiBonMagn**^{-temp}**ContFunc**_o *s'éterniser, traîner en longueur,*

on peut dégager deux composantes de sens : 'long' et 'déplaisant'. On souhaite donc pouvoir accéder à la lexie en partant de grandes catégories de type « temps » et « caractère négatif ».

Pour mener à bien cette organisation sémantique, nous avons procédé en deux étapes. La première consiste à regrouper des fonctions lexicales sémantiquement proches. Par exemple, les fonctions **Fact**, **Real** et **Labreal**, habituellement catégorisées parmi les « verbes de réalisation », comportent toutes, pour les lexies désignant un artefact, le sens de réalisation de la fonction dévolue à l'artefact. Ce qui les différencie est la présence ou l'absence des actants du mot-clé et leurs positions syntaxiques. Ainsi, pour la lexie *charrette*, dont la forme propositionnelle est :

'charrette : utilisée par l'individu X pour transporter Y et tirée par l'animal Z',

on trouve les fonctions lexicales suivantes :

Vulgarisation	Fonction lexicale	Valeur	Régime de la valeur
/~Transporter Y dans ~~/	Labreal ₁₂	charrier, charroyer	[N=Y]
/~ transporte Y~/	Fact ₂	transporter	[N=Y]
/~ fonctionne~/	Fact _o	rouler	
/~Utiliser ~~/	Real ₁	conduire, mener	[ART ~]

Tableau 1 - Ensemble de fonctions lexicales sémantiquement proches liées à la lexie *charrette*

L'idée est de pouvoir accéder à toutes ces relations par une catégorie de type « réalisation/utilisation » qui les rassemble toutes. Dans une seconde étape, nous attribuerons donc, à chaque fonction lexicale une méta-fonction qui permettra de regrouper les fonctions lexicales sémantiquement proches sous une même dénomination et d'y accéder de façon

globale. À partir des regroupements de fonctions lexicales, plusieurs catégories sémantiques ont pu émerger dont les principales sont : 'réalisation/utilisation', 'équivalence', 'localisation', 'manière', 'aspect/temporalité', 'qualification', 'appréciation', 'nombre', etc.

3.2. *Perspective « partie du discours »*

La perspective « partie du discours » permet de classer les relations lexicales en fonction des parties du discours de la valeur. Dans le DiCo, on envisage les parties du discours non seulement en prenant en compte la nature de l'unité lexicale mais également son emploi en discours. Par ailleurs, les valeurs de collocations ne se limitent pas à des unités lexicales monolexémiques mais sont parfois des locutions. Ainsi, il est courant de rencontrer des valeurs mêlant adjectifs et syntagmes adjectivaux à emploi adjectival :

/*[X] qui éprouve ~*/ *haine* **A₁** *haineux, porté [à la haine],*

ou encore, adverbess et syntagme prépositionnel employé comme adverbe :

/*[X faire qqch.] en manifestant ~*/

haine **Adv1Sympt1** *avec [haine], haineusement*

Cette vision plus large des parties du discours se révèle très pertinente dès lors que l'on utilise la base de données comme outil d'aide à la rédaction.

3.3. *Perspective « communicationnelle »*

La perspective communicationnelle vise à aider un utilisateur à choisir une unité lexicale par rapport à la structure d'une phrase à produire. La place relative de l'expression d'un prédicat et de ses actants dans la phrase contribue à ce que l'on appelle la structure communicative de l'énoncé — notamment pour ce qui est de la dichotomie thème/rhème (Mel'čuk, 2001). Par exemple, pour exprimer la relation entre la lexie *flèche* et une lexie désignant la cible atteinte par une flèche, on peut recourir à plusieurs paraphrases. Si l'on souhaite thématiser, c'est-à-dire, mettre en avant la flèche, on dira par exemple : *la flèche atteint/touche la cible*. La flèche est thème et la cible, rhème. Si, au contraire, on désire mettre la lexie désignant la cible en thème, on dira, par exemple : *la cible reçoit une flèche*. Dans le cadre de la Lexicologie Explicative et Combinatoire, on utilise, pour réaliser ces manipulations paraphrastiques, des règles de paraphrasage faisant appel au système des fonctions lexicales (Mel'čuk, 1988 ; Milićević, 2003). Pour illustrer à nouveau ce phénomène et son intérêt dans le processus de consultation d'une base de données lexicales, prenons l'exemple de l'unité lexicale *crainte* qui a la forme propositionnelle suivante :

'crainte : de l'individu X causée par Y'.

La perspective communicationnelle présente à l'utilisateur différentes collocations en fonction de la position syntaxique des éléments dans la phrase :

• **Si X est sujet et crainte (et éventuellement Y) est complément, on dira :**

X éprouve/ressent/nourrit de la crainte vis-à-vis de Y ; X conçoit/entretient de la crainte vis-à-vis de X

• **Si Y est sujet et crainte (et éventuellement X) est complément, on dira :**

Y cause/engendre/provoque/suscite de la crainte chez X ; Y inspire de la crainte à X

• **Si crainte est sujet du verbe et X est complément, on dira :**

La crainte dévore/habite/hante/ronge X

3.4. *Perspective « actantielle »*

Nous proposons une autre forme d'organisation des relations lexicales propres à une unité lexicale : la perspective actantielle. Elle permet d'organiser les relations lexicales en fonction

des actants concernés par la relation. Par exemple, *applaudissements#1* a la forme propositionnelle suivante :

'applaudissements#1 : de l'individu X destinés à l'individu Y pour son action Z'.

On cherche à répartir les relations lexicales en fonction des actants concernés. On distingue ici trois actants sémantiques X, Y et Z en correspondance avec trois actants syntaxiques 1, 2 et 3 (X=1, Y=2, Z=3), ce qui nous permet de répartir des relations lexicales en fonction des actants concernés, par exemple :

Premier actant : X = 1

/*Ensemble de X qui émettent ~*/ **QS1** *assemblée, audience, foule, public*

Deuxième actant : Y = 2

/*[Y] être visé par ~*/ **Oper2** *recueillir, se valoir [ART ~]*

Troisième actant : Z = 3

/*[Z] être la raison de ~*/ **Oper3** *provoquer, soulever, susciter [ART ~]*

Une telle organisation des relations lexicales selon différentes perspectives va donner lieu à des outils de consultation de la base de données DiCo et fournir un cadre de recherche qui facilite et oriente l'accès à l'information. Nous n'avons pu présenter ici que les principales perspectives d'organisation envisagées, notons que nous prenons également en compte dans la structuration des relations lexicales, des éléments comme le mode d'expression au style direct ou indirect (par exemple, la distinction entre : « *Au lit !* » et *envoyer qqn se coucher*), les onomatopées (par exemple, pour la lexie *horloge* : *faire ding-dong* ou *carillonner*) ainsi que les variations de registres de langue (par exemple, *biffeton* en tant que synonyme de *billet* sera marqué comme familier). Ces éléments feront l'objet d'autres perspectives ou sous-perspectives de classement.

Conclusion

Nous avons cherché à réaliser un mode de structuration des relations lexicales par l'exploitation d'un outil théorique de modélisation, les fonctions lexicales. Nous avons brièvement discuté les classements précédents, et proposé d'organiser les fonctions lexicales selon une approche multidimensionnelle. Ce modèle de structuration a l'avantage de prendre en compte plusieurs types d'information (partie du discours, sens, patrons syntaxiques, actants) de façon indépendante. Les descriptions attribuées à chaque fonction lexicale seront compilées dans une base de données de fonctions lexicales à partir de laquelle il sera possible de faire des requêtes générales ou très précises sur les relations lexicales. Cette base de données ainsi que la classification selon plusieurs dimensions se révéleront une mine de ressources pour la création d'applications liées à la consultation du dictionnaire ou à l'aide à la rédaction. L'organisation des fonctions lexicales va permettre le développement d'un grand nombre d'applications liées à l'utilisation du DiCo. Nous avons montré dans Jousse & Tremblay (2006) et Jousse, Tremblay & Polguère (2006) comment ce genre d'applications tirées du DiCo ou d'autres bases de données sont utiles pour l'apprentissage des collocations et la maîtrise du lexique. Nous développerons, dans la suite de notre travail, un environnement de consultation du DiCo permettant des requêtes multicritères et proposerons, à des fins pédagogiques, un certain nombre de requêtes prédéfinies pour guider les utilisateurs vers des « parcours » typiques de consultation.

Remerciements : Je tiens à remercier chaleureusement Alain Polguère, Sylvain Kahane et Lucie Barque pour la relecture de cet article.

Références

Références bibliographiques

- ALONSO RAMOS, Margarita (1993). *Las funciones léxicas en el modelo lexicográfico de I. Mel'čuk*, Tesis de doctorado, Madrid: UNED.
- ALONSO RAMOS, Margarita & TUTIN, Agnès (1996). A Classification and description of the Lexical Functions of the Explanatory Combinatorial Dictionary for the treatment of LF Combinations, in *Lexical Functions in Lexicography and Natural Language Processing*, WANNER, Léo (éd.). Amsterdam: John Benjamins.
- DE SCHRYVER, Gilles-Maurice (2003). Lexicographers' Dreams in The Electronic-Dictionary Age, *International Journal of Lexicography*, Vol.16, n°2: 143-199.
- FELLBAUM, Christiane (Ed) (1998). *Wordnet: An Electronic Lexical Database*. Cambridge: MIT Press.
- FILLMORE, Charles, JOHNSON, Christopher & PETRUCK, Myriam (2003). Background to FrameNet, *International Journal of Lexicography*, Vol. 16, n°3: 235-249.
- GRIMES, Joseph (1990). Inverse Lexical Functions, in *Meaning-text theory: Linguistics, lexicography, and implications*, STEELE, James (Ed), Ottawa: University of Ottawa, 350-64.
- JACQUET-PFAU, Christine (2002). Les dictionnaires du français sur CÉDÉROM. *International Journal of Lexicography*, vol. 15, no. 1: 98-104.
- JOUSSE, Anne-Laure & POLGUÈRE, Alain (2005). *Le DiCo et sa version DiCouèbe : Document descriptif et manuel d'utilisation*, Département de linguistique et de traduction, Université de Montréal.
- JOUSSE, Anne-Laure, POLGUÈRE, Alain & TREMBLAY, Ophélie (2006). Du dictionnaire au site lexical pour l'enseignement/apprentissage du vocabulaire. *Lexique et production verbale : vers une meilleure intégration des apprentissages lexicaux*, GROSSMANN, Francis & PLANE, Sylvie (Dir), coll. « Éducation et didactiques », Villeneuve d'Ascq, Presses Universitaires du Septentrion.
- JOUSSE, Anne-Laure & TREMBLAY, Ophélie (2006). Le dictionnaire électronique : un outil précieux pour enseigner le lexique. *Québec Français*, 141.
- MEL'ČUK, Igor (1988). Paraphrase et lexique dans la théorie linguistique Sens-Texte — Vingt ans après. *Revue internationale de lexicologie et lexicographie*, vol. 52/53, pp. 5-50/5-53.
- MEL'ČUK, Igor (2001). Communicative Organization in Natural Language. The semantic-communicative structure of sentences, coll. « Studies in Language Companion Series », Amsterdam/Philadelphia.
- MEL'ČUK, Igor, CLAS, André & POLGUÈRE, Alain (1995). *Introduction à la Lexicologie Explicative et Combinatoire*. Louvain-la-Neuve (Belgique), Duculot.
- MILIĆEVIĆ, Jasmina (2003). Étude des aspects sémantiques et syntaxiques de la paraphrase : application à la génération automatique des phrases françaises, Thèse de doctorat, Département de linguistique et de traduction, Université de Montréal.
- PRUVOST, Jean (Ed) (1995). Les dictionnaires de langue française et l'informatique. *Actes du colloque 1995 La Journée des dictionnaires*, Université de Cergy-Pontoise.
- PRUVOST, Jean (2000). *Dictionnaires et nouvelles technologies*, Paris : Presses universitaires de France.
- SELVA, Thierry, VERLINDE, Serge & BINON, Jean (2002). Le DAFLES, un nouveau dictionnaire électronique pour apprenants du français, *Actes du 10 congrès international EURALEX'2002, Copenhagen, Denmark, 13-17 August 2002* : 2 vls, vol I, 199-208.
- STEELE, James & MEYER, Ingrid (1990). Lexical Functions in an Explanatory Combinatorial Dictionary: Kinds and Definitions , in *Meaning-text theory: Linguistics, lexicography, and implications*, STEELE, James (Ed), Ottawa: University of Ottawa, 41-61.

STEINLIN, Jacques *et al.* (2004). De l'article lexicographique à la modélisation objet du dictionnaire et des liens lexicaux, *Proceedings of EURALEX' 2004, Lorient (France)* : 177-186.

ZHOLKOVSKY, Alexander & MEL'ČUK, Igor (1970). Towards a functioning Meaning-Text model of language, *Linguistics*, 57 :10-47.

ZOCK, Michaël & CARROLL John (2003). Dictionnaires électroniques. *Traitement automatique des langues*, 44, n°2 : 7-10.

Références aux sites Internet

Princeton University Cognitive Science Laboratory,
<http://wordnet.princeton.edu/>, (nd). Consulté en septembre 2006.

International Computer Science Institute of Berkeley,
<http://framenet.icsi.berkeley.edu/>, dernière mise à jour : 31 juillet 2006. Consulté en septembre 2006.

GRELEP, Université catholique de Louvain-la-Neuve,
<http://www.kuleuven.be/ilt/blf/>, dernière mise à jour : 09 septembre 2006. Consulté en septembre 2006.

OLST, Université de Montréal,
<http://olst.ling.umontreal.ca/dicouebe/>, dernière mise à jour : 07 septembre 2006. Consulté en septembre 2006.

APPROCHE COMPARÉE DES INTERACTIONS ENTRE APPRENANTS DANS LA CLASSE DE LANGUE ET DANS LES ENVIRONNEMENTS PÉDAGOGIQUES INFORMATISÉS

Bárbara RAZOLA MAYOR

barbara_rm@yahoo.com

Lidilem

Université Stendhal

1180 avenue Centrale - Domaine Universitaire

38400 Saint Martin d'Hères

BP 25 - 38040 Grenoble Cedex 9

Abstract : This paper compares the development of language skills through synchronic CMC and through face-to-face interaction. A recent experiment exposed a group of 20 Spanish language students to two similar communicative tasks, one on-line and one in the classroom. The results show that synchronic CMC (Computer Mediated Communication) can influence certain aspects of language acquisition.

1. La problématique

De nombreux auteurs (Ellis, 1999 ; Giacobbe, 1992 ; Pekarek, 1999, entre autres) ont signalé l'importance que l'interaction orale joue dans l'apprentissage d'une langue étrangère. Les locuteurs non-natifs, non seulement pratiquent la langue cible pendant l'interaction, mais apprennent aussi à travers celle-ci (Ellis, 1999:2).

Lors des dernières décennies, ce concept d'apprentissage par l'interaction a été incorporé à la classe de langue. Diverses solutions ont été cherchées afin de multiplier les opportunités d'interaction des apprenants en LE. L'une d'entre elles est le travail en dyades ou groupes.

Plus récemment, de nouveaux outils d'apprentissage ont été introduits dans les salles de cours. Ils proviennent, pour partie, de l'apprentissage à distance. Ils enrichissent et facilitent l'enseignement en présentiel. Ces nouveaux instruments sont issus des technologies de l'information et de la communication (TIC). C'est notamment le cas des espaces de communication synchrone écrite, plus connus comme salles de « chat » ou clavardage.

Aujourd'hui, à travers les interactions en ligne, il est possible de réaliser des tâches communicatives en groupe de façon similaire à celles de la classe de langue. Cependant, il y a encore beaucoup d'aspects méconnus qui nous empêchent d'identifier de quelle façon et à quel point ces pratiques influencent le processus d'apprentissage d'une LE.

Ainsi, notre étude se centre sur les questions suivantes :

Certains aspects de l'interaction entre non-natifs, liés la distribution des tours de parole et au taux de participation des apprenants, sont-ils modifiés dans la Communication Médinée par Ordinateur (CMO) par rapport à des échanges de même type à l'oral ?

Quels effets provoque cette influence de la CMO dans les processus d'apprentissage de la LE ?

2. Cadre méthodologique et scénario pédagogique

Afin de répondre à ces questions, nous avons constitué une grille de comparaisons entre la réalisation de deux tâches communicatives du même type : l'une en ligne, à distance et l'autre dans la salle de cours, en face à face. Le groupe ciblé se compose de 20 étudiants

français du niveau B1 et B2¹; membres d'un cours d'espagnol non spécialiste de l'Université Stendhal. Chaque apprenant a réalisé quatre tâches communicatives :

deux débats : l'un à distance et l'autre en présentiel par groupes de dix étudiants chacun.

deux jeux de rôle : l'un à distance et l'autre en présentiel par groupes de cinq étudiants chacun.

L'environnement pédagogique de cette recherche est la plateforme Esprit <http://ute2.umh.ac.be/esprit/login-index.php>. Il s'agit d'une plateforme de formation à distance inscrite dans le projet FLODI (Formations en Langues Ouvertes et à Distance Interuniversitaires) de Grenoble Universités, piloté par l'Université Stendhal – Grenoble 3.

3. Analyse et discussion

Suite à la première analyse des données, il nous est possible de dégager les premiers résultats :

3.1. La structure conversationnelle du chat : les tours de parole

De la même façon que dans la conversation orale, la base de la structure conversationnelle est le tour de parole. Selon Gallardo (1993: 32-33) il s'agit d'une « case structurale qui peut ou non être remplie par une intervention » Par ailleurs, une intervention est « l'émission d'un locuteur pendant la possession du tour de parole » Il est donc possible, d'effectuer plusieurs interventions dans un même tour de parole.

Cette organisation de la conversation orale peut être rapportée à celle de la CMO synchrone écrite. De sorte que l'intervention corresponde à l'apparition d'un message dans l'espace prévu et qu'il existe la possibilité que le tour de parole coïncide avec une ou plusieurs interventions. Dans l'exemple suivant, nous observons comment l'apprenante structure son tour de parole en quatre interventions.

Jue/chat (A)

Corinne : « si usted no quiere voy a declarar la guerra ».

Corinne : « y todo el mundo va a hacer la huelga ».

Corinne : « no hara escuela durante muchos dias ».

Corinne : « vamos a jugar todo el tiempo por la calle ».

En ce qui concerne les principes d'alternance de tours, Sack *et al.* (1974) signalent deux types de mécanismes de distribution des tours de parole : la cession du tour, lorsque le locuteur cède le tour en sélectionnant le prochain locuteur ou bien, l'auto-sélectionne, lorsque le locuteur prend la parole par propre incitative. Ces deux phénomènes présents dans les interactions orales menées par nos informants, se sont manifestés également dans la communication à distance. Pourtant nous remarquons certaines différences.

Nous commencerons par examiner le premier des mécanismes de distribution de la parole, la cession du tour. Tout d'abord, il faut signaler l'utilisation des prénoms des participants dans les données écrites pour désigner le locuteur suivant par rapport aux données orales où cet emploi reste marginal. Cela peut être provoqué par la disparition des sources d'information non linguistique visuelle ou auditive. Les locuteurs utilisent donc des mécanismes plus explicites qui évitent l'ambiguïté.

Deb/chat (AB)

Justine : « pinso que es muy negativa, **francine**! ».

Francine : « no pero hay que pensar a todo eso ».

¹ Selon le *Cadre européen commun de référence pour les langues (CECRL)*, Conseil de l'Europe, 2001, www.coe.int/T/DG4/Portfolio/documents/cadrecommun.pdf

Cependant, il est important de signaler que la cession du tour n'est pas la seule fonction des prénoms. Le rôle plus abondant est celui d'accompagner les messages de support aux opinions exprimées précédemment.

Deb/chat (AB)

Corinne : « pero el entorno no esta siempre presente ».

Francine : « **exacto Corinne** ».

Quelle est la fonction de ce type de messages? Selon Gallardo, tous les tours de parole n'ont pas la même valeur dans la conversation. Certaines interventions du type « mhm, oui, ok » n'ont pas un rôle thématique. Ce sont des signes d'attention émis par le récepteur et nommés continuateurs (Gallardo, 1993:33), régulateurs (Kerbrat-Orecchioni, 1990:18) ou phatèmes en référence à la fonction phatique selon Jakobson. Leur fonction consiste à montrer que le récepteur suit et comprend le locuteur. Cette abondance de messages de soutien souvent accompagnés des prénoms a la même fonction que celle des continuateurs. De même, le facteur plus caractéristique de ces éléments est son abondance. L'extrait suivant est un exemple de la façon dont les apprenants se soutiennent mutuellement grâce à l'usage des continuateurs.

Deb/chat (AB)

Sophie : « Los amigos de la familia no pueden ser buenos modelos para el niño porque no son presentes todos los días ».

Julie : « **yo tambien !** ».

Francine : « no es la misma cosa que una persona que vive siempre con el niño ».

Aurelie : « **exactamente sophie** ».

Justine : « y pienses que todos los padres o madres son presnte todos dias en caso de separacion? ».

Julie : « **bien dicho** ».

En outre, de nombreux exemples nous montrent comment les étudiants ont désigné le locuteur suivant sans utiliser son prénom. Cela est fait de la même manière qu'ils le font dans les données orales, c'est-à-dire, le participant dirige la parole directement au locuteur suivant à propos de quelque chose dite auparavant.

Deb/chat (AB)

Francine : « si pero no estara con los otros niños y sera muy difícil saber orientarse ».

Justine : « porque dices que no estara con otros niños? ».

Cependant, cet usage n'est pas exempt de problèmes car souvent le message dirigé directement à un locuteur précis n'est pas pris en compte (c'est le cas de l'exemple précédent). De même, ce type de messages qui vise un participant concret peut être aussi interprété comme une allusion au groupe entier. Dans l'extrait suivant, Justine pose une question directement à Francine, cependant Julie, Benoît et Corinne, et finalement Francine, répondent au message.²

Deb/chat (AB)

Francine : « no es la misma cosa que una persona que vive siempre con el niño ».

Aurelie : **exactamente sophie** ».

Justine : « y pienses que todos los padres o madres son presnte todos dias en caso de separacion? ».

² Dans ce cas, même si Justine n'utilise pas une forme correcte du verbe "pensar" le contexte nous indique que c'est une question posée uniquement à Francine à propos de son intervention en 117. Il est possible donc, que les autres étudiants interprètent la direction de cette question de façon différente à cause de la forme erronée du verbe.

Julie : « bien dicho ».
Benoît : « No pero una madre puede decir algo a su nino y un padre tambien, aun que sean separados (...) ».
Corinne : « no es la misma cosa ».
Francine : « estan presentes lo maque pueden y se preocupen del nino que han empezado a criar ya ».

Malgré les problèmes créés par cet emploi, les apprenants continuent à s’en servir. Son usage est même plus grand que celui du prénom évitant ainsi de poser des problèmes au locuteur désigné. Dans le chat AB, nous remarquons cinq cas de désignation sans prénom dont deux n’ont pas été pris en compte par leurs destinataires, et deux cas de choix avec prénom.

Désignation accompagnée d'un prénom			Désignation sans prénom		
Numéro de tour	Participant	Message attendu	Numéro de tour	Participant	Message attendu
143	Justine	Oui	27	Aurélie	Oui
159	Sophie	Oui	44	Justine	Non
			119	Justine	Oui
			146	Justine	Oui
			172	Justine	Non

Tableau 1 - Cession du tour de parole

Lorsque le changement de tour et de locuteur s’effectue grâce au procédé d’auto-sélection et non par désignation du locuteur suivant, nous apercevons des prises de tour simultanées et donc des chevauchements dans les données orales. Dans l’extrait suivant, Betty et Carine interviennent en même temps et c’est finalement Betty qui prend le tour. Par ailleurs, il est difficile de savoir pourquoi Carine s’excuse comme si elle n’avait pas le droit d’intervenir à ce moment. Ce phénomène disparaît dans le chat car la nature même de ce moyen de communication empêche les chevauchements.

Deb/oral (CD)
Anne-Claire : « (...) un cuerpo eh mutilado porque eso lo vemos todos los días casi: en las películas y todo eso y : ».
Betty : « entiendo que: ».
Carine : « pero no: pardon pardon (risas) ».
Betty : « justamente xxx justamente lo que es- estás diciendo muestra lo (...) ».

Quant à la structure interne du tour de parole, Briz (1998, 213) distingue trois types de marques qui signalent les différentes parties de celle-ci : les marques régulatrices de départ, les marques de progression et les marques de clôture. Nous remarquons leur présence dans les deux types de corpus.

Jue/oral (A)
Benoît : « hola /// **entonces** tenemos que preparar las vacaciones eh **para empezar** //quizás eh debemos preguntar a todo el mundo dónde: quiere ir y qué quiere hacer durante esas vacaciones el abuelo / **por ejemplo** ».
Jue/chat (D)
Carine : « **si** estoy de acuerdo **porque** no hay nada para divertirse por ejemplo no hay parques de atracciones ni campo de futbol! me parece abusivo **no?** ».

3.2. Variations de la participation

Afin d’analyser la participation dans les tâches réalisées, nous avons calculé le nombre total de mots produits par chaque apprenant. Warschauer (1996) qui a réalisé une expérience

similaire signale un taux de participation plus équilibré entre le clavardage et la conversation orale. Nos résultats corroborent ceux de Warschauer.

Le tableau suivant correspond aux débats, en ligne et en présentiel, réalisés par le groupe AB. Il est possible d'observer comment Corinne produit, dans la tâche orale, un nombre de mots supérieur au reste du groupe. Pourtant, sa production dans le chat diminue et s'équilibre avec celle des autres participants. Parallèlement, Justine qui participe peu à la tâche orale, augmente sa production dans le débat écrit et se rapproche du niveau des autres.

	Débat oral	Débat écrit
Francine	158	334
Sophie	141	289
Corinne	546	250
Julie Anne	279	266
Aurélie	251	169
Justine	62	232
Benoît	267	254
Clara	327	

Tableau 2 - Nombre total de mots produits par chaque apprenant

Dans le but d'établir les raisons de ce rééquilibrage nous allons observer les comportements de ces deux étudiantes dans la tâche orale. En ce qui concerne Corinne, nous observons que certaines de ses interventions sont assez longues. Le trait le plus significatif de ses productions est son désaccord avec le reste du groupe. Ce point de désaccord fait que les réponses de Corinne sont constamment demandées par les autres participants. Pour illustrer cette idée, nous avons schématisé un extrait de l'interaction en indiquant uniquement le nom des locuteurs :

Deb/oral(A): 35. « **Corinne**, 36. Benoît, 37. **Corinne**, 38. Aurélie, 39. **Corinne**, 40. Julie, 41. **Corinne** »

Il est possible d'observer comment Corinne intervient juste après chaque participation des autres locuteurs. Elle réagit à tous les arguments. Cela fait augmenter son taux de participation.

En ce qui concerne Justine, nous observons que ses interventions sont déterminantes mais elles sont plus faibles en nombre et en longueur. La première intervention de Justine a lieu dans le deuxième tour de parole, c'est-à-dire juste après la première intervention du modérateur. Celle-ci est très courte, seulement onze mots. En revanche, la première intervention de Corinne dans le débat a un total de 107 mots.

Dans sa deuxième intervention, Justine ne se souvient pas du mot *prohibido* (interdit). Julie lui fournit le mot qu'elle cherche. Cela peut être qualifié de positif pour l'apprentissage car il s'agit de la construction collaborative du discours, c'est-à-dire, l'étagage (« scaffolding ») Néanmoins, Julie profite de cette opportunité pour prendre le tour de parole.

Deb/oral (AB)

Justine : « y además si el niño ve que es peligroso va a: a: a: (ríe) tener gusta a mirar a esto porque es peligroso porque es »

Julie : « prohibido »

Justine : « prohibido »

Julie : « me parece de que el hecho de que la película está prohibida da la gana a los niños de verla porque es algo prohibido y pueden sentirse: (...) »

Au troisième essai de prise de tour de parole par Justine, Corinne n'est pas d'accord avec elle et l'interrompt. Justine ne peut pas terminer sa phrase.

Deb/oral (AB)

Justine : « y que la tele eh quizás eh eh di (ríe)
difuse: la los programas violente: a horas: cuando: los
niños no están: solos y: »

Corinne : « pero: ahora no hay: siempre: las películas
violencia hay también los juegos vídeo // y: y: van / no
pueden // (...) »

Dans la quatrième et dernière intervention de Justine, nous remarquons comment Sophie et Justine rivalisent pour le tour de parole. À nouveau, Justine n'est pas capable de se faire entendre et c'est Sophie qui prend la parole.

Deb/oral (AB)

Corinne : « diciembre eh había muchos muertos en la
televisión pero: eh no van a ir a la escuela sin conocer
la actualidad y: »

Sophie : « pueden »

Justine : « pero no tienen que ver las imagen »

Corinne : « sí pero no son imagen eh »

Sophie : « pueden pueden escuchar a la radio por ejemplo
/ porque las imágenes son muy violentas no piensas que
hay que: preservarles de la violencia / visual? »

Ainsi, des auteurs qui ont étudié l'effet de la CMO synchrone sur la participation ont relevé des écarts moins marqués entre les niveaux de participation des apprenants, le phénomène étant attribué au fait que les étudiants timides se sentent plus sûrs à distance. Néanmoins, au vu de nos résultats, nous pouvons déduire que la timidité n'est pas le seul facteur qui provoque l'équilibre dans la participation. Le système de distribution des tours de parole joue un rôle important dans ce domaine.

4. En guise de conclusion

Les données analysées plus haut mettent en évidence que la CMO synchrone modifie les mécanismes qui guident la conversation orale. Cette variation peut avoir une incidence directe sur l'acquisition de la LE. Ci-après, nous récapitulerons les résultats obtenus afin d'esquisser les effets possibles de l'emploi des conversations écrites sur la pratique de l'oral chez les apprenants de LE.

Tout d'abord, nous avons pu observer comment la CMO synchrone a des conséquences directes sur la participation. L'inefficacité de certains mécanismes de distribution de la parole propres à la communication orale (chevauchement, superpositions de tour, etc.) permet aux participants une plus grande liberté d'intervention dans la conversation écrite. Puisque les participants peuvent contribuer au moment qu'ils le souhaitent, le taux de participation en est plus équilibré.

Nous avons examiné également l'influence de la CMO synchrone lors de mécanismes d'allocation du tour. Comme nous avons pu apprécier précédemment, l'apprenant doit savoir affiner son message de sélection du locuteur dans le *chat* car les mécanismes utilisés à l'oral restent ambigus et peuvent provoquer des problèmes de compréhension. Ce résultat est une preuve de l'exigence de précision et d'explicitation demandée par la CMO synchrone, fait qui doit être exploité pour développer l'apprentissage de l'expression en LE.

Parallèlement, les continuateurs ou régulateurs présentent aussi cette influence des *chats*. Nous avons remarqué comment les apprenants les utilisent sous forme de soutien aux opinions des autres participants. En somme, ces continuateurs présentent les caractéristiques suivantes : ils incluent souvent le prénom du participant à qui ils sont dirigés, ils expriment l'accord avec le locuteur soutenu d'une façon très catégorique et explicite et, finalement, ils

sont abondants. Etant donné ces trois facteurs, nous pouvons déduire des possibles effets positifs dans la motivation à cause du constant appui proportionné mutuellement.

Quant à la structure interne du tour de parole, comme nous l'avons signalé, celle-ci ne semble pas être modifiée. Les apprenants continuent à utiliser des marques régulatrices de départ, de progression et de clôture. Le fait de les trouver sur la tâche écrite constitue un autre élément positif pour la pratique de la conversation en LE à travers la CMO synchrone.

L'hypothèse, enfin, d'une influence bénéfique de la CMO synchrone sur la pratique orale n'est pas à exclure. Il reste aussi à chercher des applications concrètes dans la classe afin de renforcer leurs effets bénéfiques.

Références bibliographiques

- BRIZ GÓMEZ, Antonio (1998). *El español coloquial en la conversación*. Ariel.
- ELLIS, Rod (1999). *Learning a second language through interaction*. John Benjamín Publishing Company.
- GALLARDO-PAÚLS, Beatriz (1993). *Lingüística perceptiva y conversación : secuencias*. Universidad de Valencia.
- GIACOBBE, Jorge (1992). *Acquisition d'une langue étrangère : cognition et interaction*. CNRS édition, Collection sciences du langage.
- KERBRAT-ORECCHIONI, Catherine (1990). *Les interactions verbales, I*. Paris : Armand colin.
- PEKAREK, Simona (1999). *Leçon de conversation : dynamiques de l'interaction et acquisition de compétences discursives en classe de langue seconde*. Fribourg édition universitaires.
- SACKS, Harvey, SCHEGLOFF, Emanuel A. & JEFFERSON, Gail (1974). Organization of turn-taking in conversation, *Language*, 50 : 696-735.
- WARSCHAUER, Mark (1996). Comparing face-to-face and electronic discussion in the second language classroom, *CALICO Journal*, 13(2) : 7-26.

CALCUL DYNAMIQUE DU SENS D'UN ADJECTIF : DE L'ANTE À LA POSTPOSITION, UN *MÉCHANT* CHANGEMENT DE SENS.

Fabienne VENANT

fabienne.venant@paris4.sorbonne.fr

LaLic, ISHA, Maison de la recherche

28 rue Serpente

75006 Paris

Abstract : This paper presents a new kind of model for meaning construction within the framework of continuous mathematics. Language is considered a morphodynamic system which follows the basic principles of Gestalt theory. The validity of this model is tested by implementing it. An analysis of the French adjectives *sec*, *curieux* and *méchant*, shows that this dynamic method for meaning computation can take into account the different factors of adjectival polysemy.

Introduction

Bien qu'étant au cœur de la plupart de nos activités quotidiennes, les processus qui président à la construction du sens restent aujourd'hui encore très mal connus. Comment des mots comme *vue*, *vie* ou *volant*, plus ou moins inertes quand ils sont isolés, prennent-ils vie au sein d'un slogan comme « Au volant, la vue c'est la vie » ? Autrement dit, comment se fait l'alchimie subtile qui, d'une succession de signes au sens plus ou moins précis, donne naissance à un contenu informatif ou émotionnel fort ? Pourquoi certaines phrases sont-elles ambiguës et d'autres non ? L'ambiguïté a-t-elle une pertinence cachée ou n'est-elle qu'un artefact à la vie dure ? Toutes ces questions sont à la base du travail présenté ici. Il s'agit de cerner un peu mieux les processus de construction du sens et de les modéliser. Or, toute tentative d'approche de la construction du sens se heurte au fait que la plupart des mots prennent des sens différents, mais apparentés, selon l'énoncé dans lequel ils sont insérés. Ce phénomène qu'on appelle polysémie est naturel, et il ne pose aucun problème dans une conversation courante. De fait, pouvoir jouer sur plusieurs sens constitue bien plus qu'une source d'amusement et de calembours. Richesse, souplesse, efficacité, créativité de la langue reposent sur ces glissements de sens permanents. Le modèle utilisé ici vise donc à rendre compte de la centralité de la polysémie dans les mécanismes de construction du sens.

1. Modèle

Donner une place centrale à la polysémie dans les processus de constructions du sens, ne pas la reléguer à un choix *a posteriori* dans une liste de sens pré-établie, c'est admettre que la construction du sens est un processus dynamique : les sens des différentes unités (lexicales, syntaxiques, grammaticales...) présentes dans la phrase s'influencent mutuellement, se spécifient progressivement tout en construisant un sens pour l'énoncé global. Le modèle utilisé ici est celui proposé par Victorri (Victorri & Fuchs, 1996). Il s'inscrit dans le cadre théorique des systèmes dynamiques. Le principe est le suivant : on associe à chaque mot polysémique un espace sémantique dans lequel se déploient ses différents sens. Les autres mots présents dans l'énoncé définissent une fonction potentielle et ce sont les bassins de cette fonction potentielle qui permettent de déterminer le sens pris par le mot étudié, dans l'énoncé considéré.

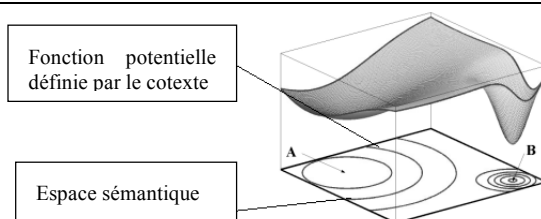


Figure 1 - Modèle (Victorri & Fuchs)

Le travail présenté ici vise à questionner ce modèle. Pour tester sa validité, nous l'avons implémenté et utilisé dans une méthode de calcul automatique du sens d'un adjectif en présence d'un nom donné.

2. Espaces sémantiques

La première étape dans l'implémentation du modèle consiste à associer un espace sémantique à chaque unité polysémique. Ploux et Victorri (1998) ont mis au point une méthode de construction automatique de ces espaces sémantiques qui repose sur l'analyse d'un graphe de synonymie (Dictionnaire Electronique des Synonymes, DES : www.crisco.unicaen.fr). Nous illustrons ici cette méthode à l'aide de l'adjectif *méchant*.

2.1. La sémantique de *méchant*

On peut distinguer pour *méchant* les principaux sens suivants (les exemples sont tirés du TLF¹ et du Petit Robert) :

- (1) Qui est de mauvaise qualité : *un méchant costume*
- (2) sans valeur, sans compétence : *un méchant billet, un méchant avocat*
- (3) qui ne remplit pas correctement sa fonction : *une méchante mémoire*
- (4) Dangereux ou désagréable, qui provoque des désagréments : *une méchante affaire*
- (5) Remarquable, extraordinaire : *une méchante voiture*
- (6) qui provoque délibérément la souffrance d'autrui : *chien méchant, méchante fée*
- (7) qui témoigne d'un désir de provoquer la souffrance d'autrui : *un regard méchant*

On voit donc se dégager dans la sémantique de *méchant*, d'une part des sens intensifs (ou désintensifs) très généraux (sens 1 à 5), c'est-à-dire qu'ils peuvent s'employer avec la plupart des noms, essentiellement en antéposition, d'autre part des sens plus riches sémantiquement (sens 6 et 7.) Le sens 6 qualifie plutôt des comportements ou des actes attestés. Le sens 7, plus psychologique, qualifie des caractères ou des intentions. On attend de l'espace sémantique que, bien que construit de façon tout à fait automatique, il rende compte de cette organisation sémantique.

2.2. Construction de l'espace sémantique

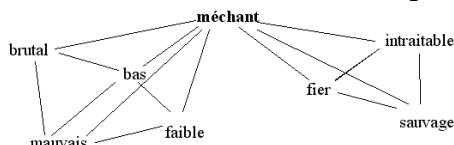


Figure 2 - un extrait du graphe de synonymie de *méchant*.

¹ Trésor de la Langue Française Informatisé (<http://atilf.atilf.fr/tlf.htm>)

Le DES fournit le graphe de synonymie de *méchant* : les sommets du graphe sont *méchant* et tous ses synonymes, il y a un lien entre deux de ces adjectifs lorsque le DES indique un renvoi synonymique. La Figure 2 montre un extrait du graphe de synonymie de *méchant*. Un synonyme ne suffit pas en général pour définir un sens du mot étudié. On voit sur la Figure 2 par exemple que *bas* est à la fois synonyme de *brutal* et de *faible*, qui correspondent à deux sens différents de *méchant*. Ploux et Victorri ont donc eu l'idée d'utiliser les cliques du graphe. Une clique est un ensemble de sommets deux à deux synonymes le plus grand possible. Le graphe de la Figure 2 présente ainsi 3 cliques : *<bas ; brutal ; mauvais ; méchant>*, *<bas ; faible ; mauvais ; méchant>* et *<fier ; intraitable ; méchant ; sauvage>*. On va considérer qu'une clique correspond, en première approximation, à une nuance de sens possible pour le mot considéré. Les cliques seront donc les points de l'espace sémantique que l'on cherche à construire. Nous ne détaillons pas ici la technique de construction de l'espace sémantique (cf. Ploux & Victorri, 1998 ou Venant, 2006). Disons simplement que l'espace sémantique est l'espace euclidien engendré par *méchant* et tous ses synonymes. A chaque adjectif correspond un axe de l'espace. A chaque clique du graphe correspond un point de cet espace dont les coordonnées dépendent des synonymes qu'elle contient. Cet espace est muni de la distance du chi2 (Bouroche & Saporta, 1980), de façon à rendre compte des proximités sémantiques réelles entre les différents sens du mot étudié. On utilise une Analyse en Composantes Principales pour obtenir une visualisation en deux ou trois dimensions de cet espace. La Figure 3 présente la visualisation de l'espace sémantique associé à *méchant* selon les deux premiers axes de l'analyse en composantes principales.

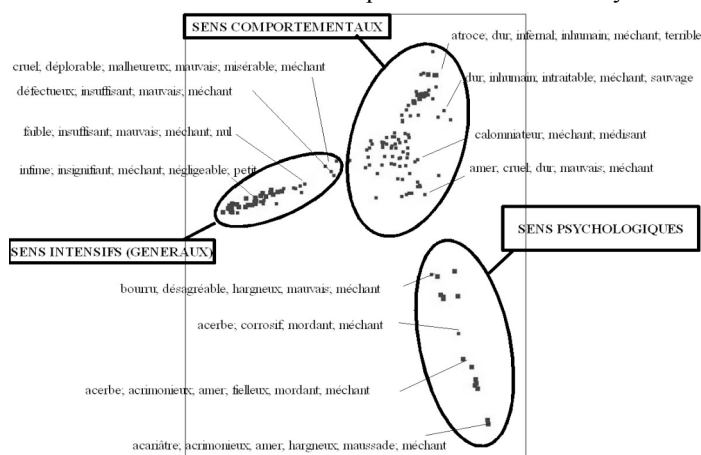


Figure 3 - Espace sémantique associé à *méchant* axes 1 et 2

On voit que l'espace construit automatiquement rend correctement compte de la sémantique de l'adjectif *méchant*. En effet, on peut constater que les sens de *méchant* se répartissent en trois zones. Dans la zone 1, en haut à gauche, on trouve les sens intensifs. La partie droite de l'espace sémantique organise les sens les plus spécifiques de *méchant*. Dans la zone 2, en haut à droite, on trouve les cliques correspondant aux sens s'appliquant surtout à des personnes et à leurs actes. Dans la zone 3, en bas à droite on trouve regroupées les cliques correspondant à des sens psychologiques de *méchant* (sens 7.)

3. Polysémie adjectivale

Nous avons dans un premier temps restreint notre étude aux adjectifs qualificatifs du français. Ils constituent en effet un champ d'expérimentation idéal car ils sont abondamment décrits d'un point de vue linguistique (Faucher, 1971 ; Goes, 1999 ; Honeste, 2001) tout en étant encore assez peu explorés d'un point de vue informatique (Bouillon & Viegas, 1999).

Nous avons ainsi pu nous faire une idée très précise des phénomènes que nous étions susceptibles de rencontrer au cours de notre étude. La polysémie des adjectifs est ainsi liée aux deux grandes caractéristiques fondamentales du sémantisme adjectival :

- une grande extension, parfois si grande qu'elle provoque ce que Goes appelle la désémantisation, au point qu'il est courant « *d'avoir une interprétation qui est si largement dépendante du nom qu'il est difficile de désintriequer ce qui dans la valeur globale obtenue vient de l'adjectif et ce qui vient du nom.* » (Goes, 1999).
- le besoin de se rattacher à un nom pour prendre tout leur sens (incomplétude référentielle chez Guillaume). Le comportement sémantique de l'adjectif est donc très dépendant de celui du nom qu'il accompagne. Il s'agit en fait d'une influence mutuelle.

Enfin, une étude du comportement adjectival ne peut ignorer les changements de sens lors du passage à l'antéposition. Ils sont complexes, non systématiques et difficiles à expliquer de façon globale. On retiendra que certains auteurs, comme Goes, proposent des explications à plusieurs facettes. Les facteurs les plus efficaces semblent être l'extension de l'adjectif d'une part, et le sémantisme du nom d'autre part.

Nous attendons du modèle qu'il soit capable de rendre compte de ces caractéristiques sémantiques des adjectifs. Avant de lancer une étude à grande échelle, nous avons voulu étudier la plausibilité du système par une étude en profondeur des adjectifs *sec*, *curieux* et *méchant*. L'étude de l'adjectif *sec* (Venant, 2004) nous a permis de vérifier que le système était capable de prendre en compte l'influence du nom recteur dans le calcul du sens d'un adjectif. Testé sur les 20 noms les plus fréquemment employés avec *sec* dans le corpus Frantext Catégorisé (20^{ème} siècle hors poésie), le système a déterminé le sens correct de *sec* dans 79% des cas. Ces résultats très prometteurs nous ont incité à continuer. Une première étude menée sur l'adjectif *curieux* (François *et al.*, 2002 commenté dans Venant, 2006) avait montré que le système était capable de rendre compte des changements de sens entre antéposition et postposition. Cette étude avait d'autre part mis en évidence le rôle de l'extension de l'adjectif dans sa prédisposition plus ou moins grande à l'antéposition. Nous poursuivons ici l'investigation de ce phénomène sur l'adjectif *méchant* dont le sémantisme en antéposition est plus complexe que celui de *curieux*. *Méchant* possède en effet une très grande extension (il peut s'appliquer à n'importe quoi, de la table au costume en passant par l'avocat, la fée, ou la pendule). En antéposition, dans ses emplois intensifs (et désintensifs), il est sujet au phénomène de désémantisation décrit par Goes. Enfin, les changements de sens entre antéposition et postposition ne sont pas systématiques. On trouve en effet des cas où le changement de sens est flagrant et d'autres où le sens de l'adjectif est le même qu'il soit placé avant ou après le nom. Ainsi s'il est préférable de ne pas engager *un méchant avocat* pour se défendre à un procès, *un avocat méchant* peut au contraire se montrer redoutable. En revanche, on redoutera de la même façon de recevoir *un coup méchant* ou *un méchant coup*.

4. Influence du nom recteur

On rend compte ici de l'aspect dynamique de la construction du sens, c'est à dire le fait que le sens du nom influence celui de l'adjectif en même temps que celui de l'adjectif influence celui du nom. On va donc procéder à une double désambiguïsation. La désambiguïsation du nom est réalisée en calculant sa classe de sélection distributionnelle (CSD) relativement à l'adjectif qui le modifie, en l'occurrence l'adjectif *méchant* (pour la définition et la méthode de construction automatique des CSD voir Jacquet & Venant, 2005). Le principe est le suivant : on cherche dans un corpus à rapprocher le nom d'autres noms qui influencent de la même manière le sens de *méchant*. Pour ce faire, on rapproche les noms partageant les mêmes contextes lexico-syntaxiques (un contexte lexico-syntaxique est par exemple « être recteur de *méchant* en tant qu'épithète antéposée » ou encore « être complément d'objet du

verbe *aimer* »). Pour obtenir automatiquement ces classes, nous avons utilisé une technique d'analyse distributionnelle sur un gros corpus analysé par l'analyseur syntaxique SYNTEX. Le principe du calcul est le suivant. À chaque unité lexicale on associe sa *fiche distributionnelle*, constituée par la fréquence relative de ses différents contextes syntactico-lexicaux. Pour obtenir les classes de sélection distributionnelle associées à un nom donné, en tant que recteur d'un adjectif donné, on établit d'abord la liste des noms qui ont été rencontrés comme recteur de cet adjectif, puis on construit automatiquement des regroupements au sein de cette liste, en calculant une distance entre ces noms à partir de leur fiche distributionnelle. Ce qui nous intéresse particulièrement ici, c'est que la classe d'un nom varie en fonction de l'adjectif étudié, et c'est cette classe qui permet ensuite de désambiguïser l'adjectif, et de lui assigner le sens correct en présence du nom considéré.

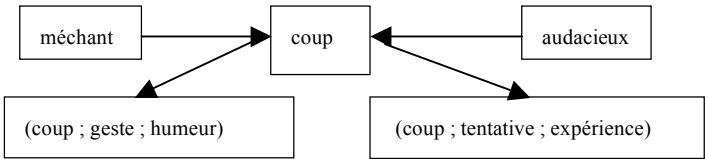


Figure 4 - La CSD associée au nom dépend de l'adjectif qui le modifie

Deux classes distributionnelles [ANTE vs. POST] sont calculées pour chaque mot, l'une à partir des fréquences en antéposition, l'autre à partir des fréquences en postposition. On mesure ensuite l'influence du nom recteur sur la sémantique de *méchant* en associant à chacune de ces classes une fonction potentielle définie sur l'espace sémantique. Rappelons qu'à chaque point de l'espace sémantique correspond une clique du graphe de synonymie de *méchant*. La valeur de la fonction en chaque point dépend des fréquences de cooccurrence (en antéposition pour une classe [ANTE], en postposition pour une classe [POST]) de chacun des noms constituant la classe avec chacun des adjectifs constituant la clique (pour le détail des calculs voir Venant, 2006). La Figure 5 montre les fonctions potentielles associées aux CSD du nom *bête*. Les bassins de ces fonctions indiquent la zone de l'espace sémantique dans lequel le nom contraint l'adjectif à prendre son sens.

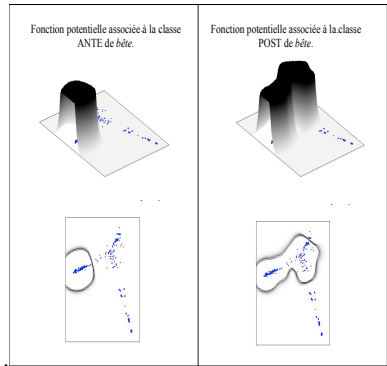


Figure 5 - Fonctions potentielles associées aux classes distributionnelles de *bête* en tant que recteur de *méchant*.

5. Désambiguïstation de l'adjectif

Il s'agit de déterminer le sens que prend l'adjectif *méchant* en présence d'un nom donné. Par sens, nous entendons ici la zone (zone 1 - sens intensifs, zone 2 - sens comportementaux ou zone 3 - sens psychologiques) de l'espace sémantique correspondant au sens de *méchant* dans le syntagme étudié. Notons que ces zones ont été définies manuellement. Le système calcule

une fonction potentielle pour chacune des zones. En chaque point, la valeur de la fonction dépend de l'appartenance ou non de la clique à la zone considérée. (Figure 6)

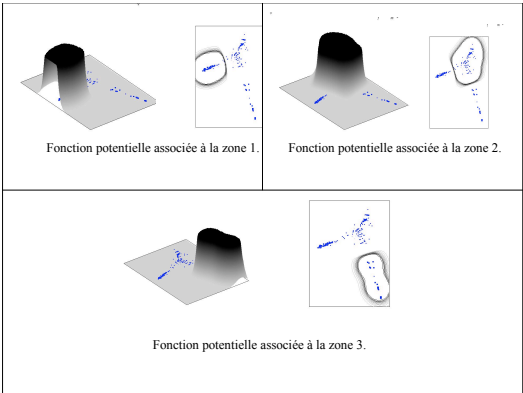


Figure 6 - Fonctions potentielles associées aux zones de sens de *méchant*

La méthode de désambiguïsation est la suivante : étant donné un syntagme *méchant* + N ou N + *méchant*, le système détermine la CSD [ANTE] ou la CSD [POST] associée à N en tant que recteur de *méchant*, il calcule ensuite la fonction potentielle associée à cette CSD et la compare à chacune des fonctions potentielles associées aux zones de sens de l'espace sémantique de *méchant*. Il calcule ainsi un degré d'affinité entre le nom N et chacune des zones de sens. Là où les zones obtenant le plus fort degré d'affinité est sélectionnée pour désambiguïser *méchant* dans le syntagme étudié. Les calculs ont été menés pour les 40 noms les plus fréquemment utilisés comme recteur de *méchant* dans le corpus Frantext Catégorisé. Le tableau 1 présente quelques exemples :

ANTE			POST	
Sens intensifs	76%	cheval	Sens intensifs	69%
Sens comportementaux	24%			
Sens intensifs	72%	esclave	Sens comportementaux	51%
Sens comportementaux	28%		Sens intensifs	28%
			Sens psychologiques	21%
Sens comportementaux	100%	regard	Sens comportementaux	100%
Sens intensifs	93%	homme	Sens intensifs	100%
Sens intensifs	99%	loup	Sens comportementaux	53%
			Sens intensifs	47%

Tableau 1 - Exemples de résultats

Les résultats que nous obtenons rendent compte des changements de sens entre antéposition et postposition. Ce sont les sens intensifs qui obtiennent les scores les plus élevés dans quasiment tous les cas en antéposition. Cela illustre l'idée qu'on trouve en antéposition les valeurs de sens qui ont la plus grande extension. Ici c'est clairement la valeur intensive qui a la plus grande extension. Elle peut s'appliquer à n'importe quoi alors que les deux autres valeurs ne s'appliquent qu'à des noms animés ou considérés comme tels (on trouve dans les contes de fées *des bois méchants* et autres *chemins méchants* qui font du mal aux gens qui s'aventurent en leur sein). Selon Honeste (2001), on peut considérer qu'en antéposition, l'adjectif est conceptualisé avant le nom. Il présente d'abord un signifié informe car non encore configuré par le support. Le nom vient dans un deuxième temps fournir le support attendu et la notion qu'il exprime s'associe avec celle de l'adjectif sans être modifiée. Si en antéposition les sens intensifs sont omniprésents, il y a des noms pour lesquels ils ne s'imposent pas forcément. *Cheval, couleur, eau, espèce, farce, matin, mot, nature, parole,*

taureau, terre donnent à *méchant* en antéposition tantôt une valeur générale, tantôt une valeur comportementale. Le contexte permet souvent de trancher entre les deux valeurs mais ce n'est pas toujours très clair. On illustre ici le phénomène sur le nom *cheval*.

A. *Sur une route quelconque, à travers la plaine immense et quelconque, au trot de deux méchants chevaux, nous allons* ». [Barrès M]

B. « *Non, non, je veux voir maman; ces méchants chevaux ont emporté maman* ». [Comtesse de Ségur]

Il est difficile de déterminer avec certitude quelle est la valeur sémantique de *méchant* dans ces énoncés. Dans l'énoncé A, la répétition de *quelconque* nous fait pencher pour une valeur générale mais au fond le doute est permis. Dans l'énoncé B, on penche pour une valeur comportementale, tant l'enfant a l'air triste, mais cette valeur n'est pas fixée. Selon les noms concernés, il s'agit soit d'une indétermination, soit d'une ambiguïté, qui de toute façon disparaît au passage à la postposition. Faucher (1971) propose une explication à ce phénomène : l'adjectif postposé garde son indépendance sémantique vis à vis du nom, ce qui lui confère une valeur distinctive. Il restreint l'extension du nom. La propriété convoquée s'applique alors seulement aux éléments de la classe convoqués dans le discours et ne peut pas être reprise en fonction attributive pour la classe entière. Nos résultats montrent que notre système est capable de rendre compte de ce phénomène. *Cheval, couleur, dent, eau, espèce, farce, maison, mot, nature, parole, part, société, taureau, terre* acceptent aussi bien une valeur générale que comportementale en antéposition, mais ne sélectionnent que la valeur comportementale ou psychologique en postposition. Pour les noms *bête, bois, bruit, chemin, chose, corps, rire* et *voix* le changement de sens est encore plus radical puisqu'en antéposition *méchant* est exclusivement général alors qu'en postposition il devient comportemental ou psychologique. Le système est capable aussi de repérer les noms pour lesquels on ne repère pas de changement de sens lors du passage de l'antéposition à la postposition. C'est le cas ici de *coup, part, regard* et *vérité*.

Afin de mesurer l'apport des classes distributionnelles, et donc de la désambiguïsation simultanée du nom et de l'adjectif, ces résultats ont été comparés à ceux obtenus, pour les mêmes noms et sur la même tâche, sans utiliser les classes distributionnelles, c'est-à-dire en associant une fonction potentielle au nom seul, selon ses fréquences d'emplois avec les différents adjectifs. Il s'avère que l'utilisation des classes distributionnelles ne fait perdre aucun des résultats positifs qui apparaissent avec le nom seul. Elle apporte au contraire quelques nuances de sens (comme par exemple le fait qu'un *méchant esclave* l'est au sens générique, en tant qu'archétype social, avec la sélection des sens intensifs, mais aussi d'un point de vue comportemental, son écart au type de l'esclave modèle pouvant se traduire par un comportement violent). L'utilisation des classes distributionnelles permet aussi d'obtenir des résultats corrects là où l'utilisation du nom seul génère une erreur ou ne donne pas de résultat du tout. Les erreurs rencontrées sur des noms comme *homme* ou *enfant* (calcul d'une valeur générale en postposition) ne sont pas le fait de l'utilisation des CSD. Elles montrent clairement les limites de la synonymie comme description du sens. Le calcul d'une valeur générale en postposition repose sur les hautes fréquences de cooccurrences de ces noms avec les adjectifs *maigre, faible, pauvre* et *petit*. Or la synonymie entre *méchant* et ces adjectifs n'est plus valable dans le contexte des noms considérés ici, en présence desquels *méchant* se colore plutôt d'une valeur comportementale ou psychologique. Le calcul d'un sens intensif en postposition pour *cheval* montre par ailleurs que ces relations de synonymie, en plus d'être partielles, ne sont valables qu'en antéposition. Nous travaillons à une méthode d'exploration du graphe adjectival global, qui permettra de repérer automatiquement les sens intensifs d'un adjectif, ceux qui ne sont valables qu'en antéposition.

Conclusion

Les analyses détaillées que nous avons menées montrent que les outils informatiques développés sont très prometteurs. Ils peuvent déjà être appliqués tels quels à la désambiguïsation automatique de n'importe quel couple nom + adjectif. Les différentes étapes dans la réalisation et l'analyse des résultats obtenus à chaque pas ont mis au jour différents problèmes, que nous devons résoudre pour aller plus avant. L'utilisation des cliques du graphe de synonymie s'est avérée fort judicieuse, tant pour la construction des espaces sémantiques que pour le calcul du sens proprement dit. Le recouvrement de l'espace sémantique par les cliques contrebalance le fait que la relation de synonymie est une relation partielle, non transitive et peut dépendre de la position de l'adjectif, ce qui cause cependant encore quelques problèmes non résolus. Les outils réalisés ont prouvé que leur utilité pouvait dépasser le cadre du calcul automatique du sens et fournir des éléments de réflexion sur des phénomènes aussi subtils que les changements de sens entre antéposition et postposition. Le travail décrit ici ne constitue cependant que la première étape dans le processus de validation du modèle utilisé. Nous devons maintenant mener des évaluations massives sur un échantillon beaucoup plus large d'adjectifs ambigus, pour ensuite dépasser le cadre du syntagme, passer à l'énoncé voire au texte tout entier. Ce travail laisse cependant entrevoir tout l'intérêt de l'utilisation des mathématiques du continu en linguistique.

Références

- BOUILLON, Pierrette & VIEGAS, Evelyne (1999). The description of adjectives for natural language processing : theoretical and applied perspectives, in *Atelier thématique sur la description des adjectifs pour les traitements informatiques*, Institut d'études scientifiques de Cargèse.
- BOUROCHE, Jean-Marie & SAPORTA, Gilbert (1980). *L'analyse des données*. Paris : Presses Universitaires de France.
- FAUCHER, Edith (1971). La place de l'adjectif. Critique de la notion d'épithète, *Le Français Moderne*, 39, 119-127.
- FRANÇOIS, Jacques, VICTORRI, Bernard & MANGUIN, Jean-Luc (2002). Polysémie adjectivale et synonymie : l'éventail des sens de curieux, in *La polysémie*, Soutet, Olivier (Eds). Paris : Presses Universitaires de l'Université de la Sorbonne.
- GOES, Jan (1999). *L'adjectif. Entre nom et verbe*. Paris/Bruxelles : Duculot.
- HONESTE, Marie-Luce (2001). Approche cognitive de la fonction adjectivale, in *L'adjectif en français et à travers les langues*, François, Jacques (Eds). Caen : Presses Universitaires de Caen.
- JACQUET, Guillaume & VENANT, Fabienne (2005). Construction automatique de classes de sélection distributionnelles, in *Actes du colloque TALN'05*, Dourdan.
- PLOUX, Sabine & VICTORRI, Bernard (1998). Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes, in *TAL* 39(1)
- VENANT, Fabienne (2004). Polysémie et calcul du sens, in *Le poids des mots, Actes des 7^{ème} journées internationales d'Analyse des Données Textuelles*, Louvain la neuve.
- VENANT, Fabienne (2006). *Représentation et calcul dynamique du sens : exploration du lexique adjectival du français*. Mémoire de doctorat de l'EHESS.
<http://cyberdocs.univ-lyon2.fr/index.php/record/view/3678>
- VICTORRI, Bernard & FUCHS, Catherine (1998). *La polysémie. Construction dynamique du sens*. Paris : Hermès.

LA PLATE-FORME LINGUASTREAM

Antoine WIDLÖCHER, Frédéric BILHAUT

{awidloch,fbilhaut}@info.unicaen.fr

Laboratoire GREYC CNRS UMR 6072, Université de Caen

14032 Caen Cedex

Abstract : This article describes the LinguaStream platform, thereby also introducing various methodological principles and analysis models. These models can be applied to the elaboration of complex NLP systems by articulating heterogeneous processing tasks, for research or educational purposes. A typical application is described in detail concerning the problem of temporal discourse framing.

1. Introduction

LinguaStream¹ (Bilhaut & Widlöcher, 2006a; Widlöcher & Bilhaut, 2005; Bilhaut & Widlöcher, 2005b) a été initialement développée pour faciliter la réalisation d'**expériences sur corpus en TAL** (Traitement Automatique des Langues), ainsi que le cycle d'évaluation/ajustement qui en découle. Sans outil adapté, le coût de mise en œuvre induit par chaque nouvelle expérience devient en effet un frein considérable à l'approche expérimentale, ainsi qu'à toute application pédagogique où l'on souhaite se concentrer sur les modèles et règles linguistiques. Pour répondre à cette problématique, LinguaStream permet de mettre en œuvre des procédés non triviaux tout en requérant des compétences informatiques minimales. Plate-forme générique fondée sur le principe d'**enrichissement incrémental** des documents électroniques, et visant l'**articulation de traitements sur corpus** (Bilhaut & Widlöcher, 2005a), elle facilite la conception et l'évaluation de **chaînes de traitements** complexes, par assemblage visuel de modules d'analyse de types et de niveaux variés : morphologique, syntaxique, sémantique, discursif... Chaque palier de la chaîne de traitement se traduit par la découverte et le marquage de nouvelles informations, sur lesquelles pourront s'appuyer les analyseurs subséquents.

Un **environnement de développement intégré** (cf. figure 1) permet de construire visuellement ces chaînes de traitement, à partir d'une « palette » de composants (une cinquantaine est intégrée en standard, cet ensemble étant extensible si besoin). Certains sont spécifiquement dédiés à des traitements d'ordre linguistique, et d'autres permettent de résoudre différents problèmes liés à la gestion des documents électroniques (traitements XML en particulier). D'autres peuvent être utilisés pour effectuer des calculs sur les annotations produites par les analyseurs, pour générer des diagrammes, etc. D'autres encore permettent de **visualiser** les documents analysés et leurs annotations selon différents modes : corpus annoté, vue de type concordancier, filtrage de certains éléments textuels... Chacun dispose d'un ou plusieurs points d'entrée et/ou de sortie que l'on relie pour obtenir la chaîne voulue, celle-ci étant représentée par un graphe où les divers composants apparaissent sous forme de « boîtes » reliées entre elles. Chaque composant propose un nombre variable de paramètres permettant d'adapter son comportement. Les marquages produits par chacun sont organisés en couches indépendantes, supportant enchâssements et chevauchements. La plate-forme se base systématiquement sur les **standards XML**, et peut traiter tout fichier de ce type en préservant sa structure originelle. Il est ainsi possible d'exploiter directement des

¹ Différentes démonstrations, un manuel d'utilisation et une bibliographie sont disponibles sur le site web dédié à la plate-forme (Bilhaut & Widlöcher, 2006b).

corpus variés (XHTML, DocBook, etc.) et de ne rien perdre de leur format d'origine, et donc, par exemple, de les visualiser simplement, par les moyens classiques : un corpus XHTML restera par exemple visualisable dans un navigateur web.

La plate-forme permet ainsi d'appréhender des problèmes de natures variées, à des échelles allant du caractère au discours, en passant par les niveaux lexical, syntagmatique et phrastique, et en s'appuyant sur des traitements différents à chaque palier de l'analyse. On peut ainsi, par exemple, explorer sur corpus la distribution des éléments de tel ou tel lexique, procéder à l'analyse sémantique des expressions temporelles, mettre en place un système de résumé par extraction, analyser la structure thématique ou rhétorique du discours, etc.

LinguaStream partage avec des outils tels que GATE (Cunningham *et al.*, 2002) et NOOJ (Muller *et al.*, 2004) un certain nombre de principes. Elle s'en distingue cependant, en particulier, en privilégiant la *combinaison* de modèles théoriques variés, en favorisant les formalismes *déclaratifs*, en permettant l'élaboration *graphique* des chaînes de traitement et en autorisant l'utilisation de tout type de corpus XML. L'évocation des principes fondamentaux de la plate-forme éclairera ces différents aspects.

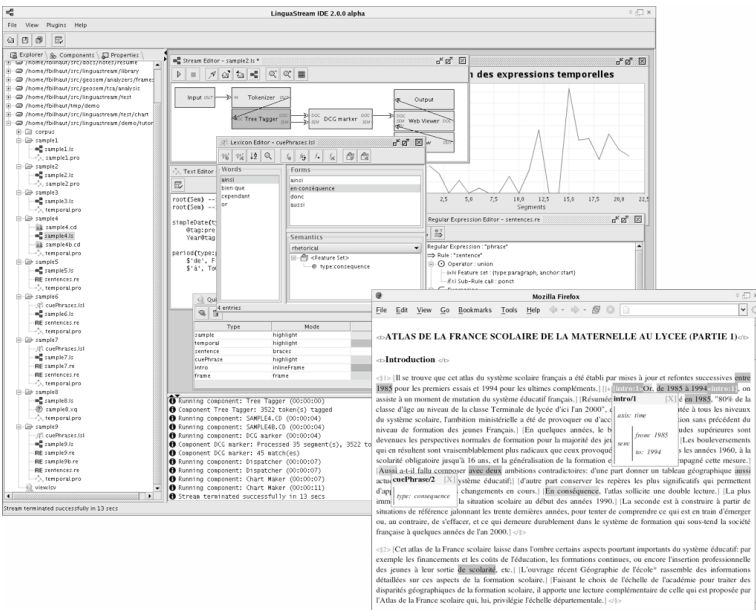


Figure 1 - Environnement de LinguaStream et visualisation avec Firefox

2. La plate-forme LinguaStream

2.1. Principes fondamentaux

En premier lieu, la plate-forme recourt systématiquement à des **représentations déclaratives** pour spécifier les différents traitements, ainsi que leur enchaînement sous forme de graphe. Les différents formalismes disponibles permettent ainsi de transcrire directement l'expertise linguistique à mettre en œuvre, l'appareil procédural qui en résulte étant pris en charge par la plate-forme. Les règles données ont donc une valeur tant **descriptive**, en tant que représentations formelles d'un phénomène linguistique, que **prescriptive**, en tant qu'instructions de traitement fournies à un processus informatique.

De plus, la plate-forme exploite la **complémentarité des modèles d'analyse**, plutôt que de privilégier un hypothétique modèle « omnipotent » capable d'exprimer efficacement tout type de contrainte. Nous faisons en effet l'hypothèse qu'un analyseur complexe doit adopter

successivement plusieurs regards sur le même matériau linguistique, auxquels répondront des formalismes distincts. On pourra combiner, au sein d'un même traitement, des expressions régulières au niveau morphologique, une grammaire locale au niveau syntagmatique, un transducteur au niveau phrastique et une grammaire CDML (cf. *infra*) au niveau discursif. Si certains modèles d'analyse s'avèrent plus adaptés à certains types de problèmes et à l'écriture de règles à certains niveaux de granularité, l'utilisateur est cependant entièrement libre de ce choix. L'interopérabilité des différents modèles d'analyse proposés est garantie par l'usage d'une **représentation unifiée** des marquages et des annotations. Ces dernières sont uniformément représentées par des **structures de traits**, modèle communément utilisé en TAL et en linguistique, et permettant de représenter des annotations riches et structurées. Tout composant d'analyse peut produire son propre marquage en s'appuyant sur les analyses précédentes : les formalismes proposés permettent de spécifier des contraintes sur les annotations existantes par **unification**. La plate-forme favorise ainsi l'**abstraction progressive des formes de surface**. Chaque palier d'analyse pouvant accéder simultanément aux annotations produites par tous les paliers antérieurs, les analyseurs de plus haut niveau sont généralement conduits à s'abstraire progressivement du matériau textuel pour ne plus reposer que sur des représentations symboliques antérieurement calculées, comme l'illustre la figure 2.

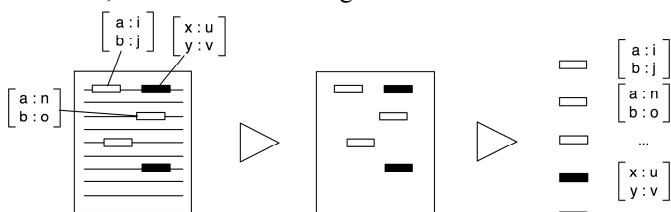


Figure 2 - Abstraction progressive des formes de surface

Un autre aspect important concerne la **variabilité du grain d'analyse** au cours du traitement. De nombreux modèles d'analyse imposent la définition d'un grain d'analyse minimal, dit « jeton » ou *token*. C'est par exemple le cas de toute grammaire ou transducteur : ces formalismes supposent l'existence d'une unité textuelle (comme le caractère ou le mot) à laquelle s'appliquent les patrons. Quand la définition de ce grain minimal est nécessaire au fonctionnement d'un composant, la plate-forme permet de spécifier **localement** le ou les types d'unités à considérer comme jetons. Toute unité préalablement délimitée peut jouer ce rôle. Il pourra s'agir du découpage habituel en mots, mais aussi de toute autre unité ayant été préalablement marquée : syntagmes, phrases, cadres du discours, etc. Le grain minimal peut donc être différent pour chaque palier de l'analyse, ce qui augmente considérablement la portée des différents modèles d'analyse utilisables dans la plate-forme. D'autre part, chaque module d'analyse spécifie les marquages antérieurs auxquels il souhaite faire référence, pouvant ainsi ne tenir compte que des marquages qu'il estime pertinents, et donc s'affranchir partiellement de la linéarité du texte. La combinaison de ces fonctionnalités permet d'adopter une **perspective d'analyse** sur le document spécifique à chaque étape d'une chaîne de traitement.

La **modularité** des chaînes de traitement favorise quant à elle la **réutilisabilité** des composants dans des contextes différents : un module d'analyse développé au sein d'une première chaîne pourra être réutilisé dans d'autres chaînes. De façon similaire, toute chaîne pourra être réutilisée en tant que constituant d'une chaîne de plus haut niveau, sous forme de « macro-composant ». Réciproquement, pour une chaîne donnée, on pourra **substituer** à un composant tout autre composant **fonctionnellement équivalent**. Par exemple, pour une sous-tâche donnée, un prototype rudimentaire pourra être remplacé *in fine* par un équivalent pleinement opérationnel. Ceci rend possible la mise en comparaison des traitements, en

soumettant ces derniers à des contextes rigoureusement identiques, condition *sine qua non* d'une comparaison pertinente.

2.2. Composants dédiés au TAL et modèles d'analyse

Nous avons évoqué quelques-uns des composants susceptibles de prendre part à une chaîne de traitement. Parmi ceux spécifiquement dédiés au TAL on peut distinguer deux familles.

La première regroupe les **analyseurs « prêts à l'emploi »**, dédiés à une tâche précise. Il s'agira par exemple de l'étiquetage morpho-syntaxique, une interface avec TreeTagger (Schmid, 1994) étant intégrée par défaut, ou syntaxique en s'appuyant sur les résultats de Syntex (Bourigault & Fabre, 2000). Ces composants sont paramétrables, mais il n'est pas possible de modifier fondamentalement leur fonctionnement.

D'autres, au contraire, proposent un **modèle d'analyse**, c'est-à-dire un formalisme de représentation de contraintes linguistiques, éventuellement associé à un modèle opératoire, par lequel l'utilisateur peut spécifier intégralement le traitement à opérer. Ils permettent d'exprimer des contraintes tant sur les formes de surface que sur les annotations insérées par les analyseurs précédents. Toutes les annotations sont représentées sous forme de structures de traits, et les contraintes sont systématiquement spécifiées par unification sur ces structures. Quelques-uns des systèmes proposés sont :

- Un système appelé EDCG (pour *Extended-DCG*), permettant d'écrire des **grammaires locales d'unification** en se basant sur la syntaxe DCG (*Definite Clause Grammars*) de Prolog. Une telle grammaire peut être décrite dans le plus pur style déclaratif, bien que les spécificités du langage logique restent accessibles aux utilisateurs expérimentés.
- Un système, nommé MRE (pour *Macro-Regular-Expressions*), permettant de décrire des patrons par **transducteurs**, s'appliquant aussi bien aux formes de surface qu'aux annotations préalablement calculées. Sa syntaxe est similaire à celle des expressions régulières communément utilisées en TAL et en linguistique sur corpus. Mais à la différence de ces dernières, ce formalisme ne s'applique pas spécifiquement aux caractères ni aux mots, et peut porter sur toute unité textuelle préalablement analysée.
- Un langage d'expression de **contraintes au niveau discursif**. Le formalisme CDML (pour *Constraint-based Discourse Modeling Language*) (Widlöcher, 2006), que nous évoquerons ci-après, permet l'exploration des organisations discursives par l'expression et la satisfaction de contraintes, pouvant être non séquentielles, exprimées à l'aide d'un ensemble de fonctions discursives primitives (présence/absence, cohérence sémantique...), et pouvant porter en particulier sur les annotations produites en amont et sur des relations entre ces dernières.
- Un système d'annotation à partir de **lexiques sémantiques**, un outil de **tokenisation** basé sur des expressions régulières (au niveau caractère), un système permettant de délimiter des objets linguistiques déduits du balisage XML du document, etc.

3. Applications pédagogiques

Ces principes fondamentaux rendent l'usage de la plate-forme pertinent dans le cadre de l'enseignement de la linguistique et du TAL. La dissimulation de l'appareil procédural au profit des formalismes permettant l'expression de règles simplifie en particulier la mise en lumière de l'expertise linguistique jugée pertinente. Les principes de modularité permettent pour leur part d'isoler un problème singulier, de nature (morphologique, syntaxique...) et de grain (mot, phrase, discours...) variables, sans perdre le bénéfice des analyses préalables éventuellement nécessaires, en considérant simplement leur apport comme une « donnée » accessible. Les différents modes de visualisation proposés permettent enfin de rendre les

phénomènes étudiés à la fois observables et « tangibles », sans qu'il ne soit nécessaire de repenser la manière dont des « vues » pourront être produites.

Un certain nombre d'expériences pédagogiques confirment cette adéquation. Que ce soit pour l'enseignement de notions linguistiques ou l'étude de leurs applications en TAL, auprès d'auditoires de spécialistes et de non spécialistes, la plate-forme a en effet déjà fait l'objet d'utilisations fructueuses, à l'université de Caen et à celle de Toulouse le Mirail, en collaboration avec des chercheurs du GREYC (Groupe de Recherches en Informatique, Image, Automatique et Instrumentation de Caen) et de l'ERSS (Equipe de Recherche en Syntaxe et Sémantique).

4. Exemple d'application en analyse du discours

Afin d'illustrer les avantages de tels principes et de tels outils, envisageons à présent un exemple de traitement pouvant être ainsi mis en œuvre. Nous retenons ici l'exemple révélateur d'un traitement en analyse de discours choisi pour la variété des « points de vue » qu'il exige. Évoquons tout d'abord les avantages que l'analyse automatique du discours peut tirer des principes proposés. Un premier apport significatif résulte de l'approche par enrichissement incrémental et par abstraction progressive des formes de surface. S'il est naturel d'opérer au niveau de ces dernières pour une analyse par exemple morphologique ou syntaxique, il va sans dire que l'analyse discursive ne peut s'accommoder de la diversité combinatoire apparaissant à ce niveau, et qu'un filtrage s'impose. En plus de la possibilité d'opérer la pure et simple *occultation* d'éléments peu pertinents pour tel ou tel besoin interprétatif, la plate-forme permet d'opérer ladite abstraction de deux manières complémentaires. En premier lieu, l'unicité du modèle de marquage et d'annotation donne à chaque étape d'analyse l'accès aux représentations symboliques produites en amont, et permet ainsi de ramener la diversité combinatoire de surface à celle des valeurs interprétées, généralement moins nombreuses. En second lieu, le principe de variabilité du grain d'analyse déjà évoqué permet d'exploiter au niveau discursif des modèles d'analyse habituellement dédiés à des niveaux de granularité inférieurs. Par exemple, des règles EDCG pourront aussi bien décrire des patrons syntagmatiques qu'une grammaire textuelle, selon le grain choisi.

Par ailleurs, la plate-forme propose des modèles d'analyse spécifiquement adaptés au niveau discursif. Le langage CDML, en particulier, s'écarte des paradigmes généralement adoptés par les autres formalismes (y compris MRE ou EDCG), qui reposent fondamentalement sur des principes de *linéarité* (on tient compte de tous les éléments successifs) et de *séquentialité* (un ordre est imposé), principes souvent inadaptés au niveau discursif. En permettant l'expression de contraintes *non séquentielles* et *non linéaires*, le formalisme CDML autorise l'expression et la détection de motifs pouvant porter sur des éléments distants dans le texte, sans faire d'hypothèse sur leur ordre, ce qui s'avère particulièrement adapté à l'analyse du discours.

Afin de donner une idée plus concrète des principes méthodologiques présentés, envisageons à présent une configuration linguistique particulière, assez représentative des problèmes posés par l'analyse discursive, en abordant le problème de l'encadrement du discours (Charolles, 1997), et plus particulièrement celui de la détection automatique des *cadres temporels*. Rappelons que ce terme désigne, selon cet auteur, des segments textuels homogènes du point de vue d'un critère d'interprétation fixé dans une expression en position détachée en début de phrase, dite *introduceur de cadre*.

La figure 3 donne un exemple de cadre temporel (tel qu'annoté par la plate-forme par le traitement indiqué ci-après) : les éléments présents dans le cadre (délimité par des pointillés) doivent, pour être interprétés, être rapportés au critère sémantique fourni par l'expression temporelle en position d'introduceur : « Depuis le milieu des années 1980 ».

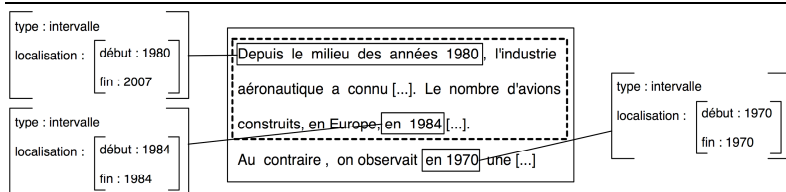


Figure 3 - Exemple de cadre temporel

L'opérationnalisation en TAL de ce modèle psycho-linguistique impose la résolution de deux problèmes principaux : la détection des introducteurs, puis l'évaluation de leur *portée*, c'est-à-dire la détermination de la borne droite du cadre introduit.

Le problème de la détection des introducteurs temporels se décline lui-même en deux sous-problèmes : l'analyse des expressions temporelles, et celle des introducteurs s'appuyant sur elles. Les principes de modularité déjà évoqués trouvent ici leur justification, puisque nous souhaiterions généralement traiter ces problèmes indépendamment. L'analyse sémantique des expressions temporelles fait l'objet d'une grammaire EDCG, exprimant des contraintes sur les résultats d'une analyse morpho-syntaxique préliminaire, et associant aux expressions reconnues une représentation de leur « sens » sous forme de structures de traits. L'extrait de grammaire suivant, évidemment simplifié à l'extrême, ne fait qu'indiquer l'esprit d'une telle grammaire. Il décrit des motifs de la forme <préposition, mois> (en janvier, ...), en s'appuyant sur l'étiquetage morpho-syntaxique réalisé au préalable (en gras), et fait remonter les valeurs adéquates (entre parenthèses) dans une structure de traits (le trait 'type' prendra la valeur 'date', et le trait 'mois' la valeur '1' pour janvier...), de telle sorte que des calculs sémantiques puissent être effectués par la suite.

```
date(type:date..mois:M) --> @tag:pre, mois(M).
date(type:date..année:A) --> ...
mois(1) --> $('janvier'.
...
```

Sur cette base, la détection des introducteurs peut être mise en place à l'aide de critères principalement positionnels. Les contraintes exprimées sont fondamentalement séquentielles : nous recherchons des zones de texte vérifiant des motifs imposant la présence, dans un ordre fixé, d'éléments immédiatement successifs. Ces règles sont donc simplement exprimables à l'aide du formalisme MRE (outre les expressions temporelles, nous exploitons ici le marquage des phrases et des connecteurs de discours) :

```
{type : phrase, anchor : start}
<introduceur>
  {type : connecteur}? {type : temporel} /as $t
</introduceur> /sem {axe : temps, valeur : $t}
", "
```

Les contraintes sur les structures de traits produites en amont (ici, en gras), ainsi que sur les formes de surface (ici, la virgule en fin de motif) permettent de délimiter l'introduceur. Nous recherchons les éléments précédés d'un début de phrase et composés d'un éventuel connecteur de discours (« mais », « cependant », etc.) et d'une expression temporelle (« En 1960 », etc.). Le reste de l'expression correspond au marquage et à l'annotation produits en sortie. L'élément reconnu aura le type « introduceur » et sera associé à l'annotation sémantique qui lui fait suite. Précisons que la variable *\$t* permet de faire « remonter » l'information contenue dans la structure de traits associée à l'expression temporelle, pour un usage ultérieur.

La détermination de la portée de l'introduceur s'avère plus problématique dans la mesure où les critères formels de clôture des cadres sont difficiles à établir. Un certain nombre d'indices ont toutefois pu être dégagés dans le cas précis des cadres temporels (Bilhaut *et al.*, 2003). La méthode utilisée ici s'appuie sur des critères énonciatifs tels que la cohésion des temps

verbaux et sur des calculs sémantiques de cohérence entre l'introducteur et les autres expressions temporelles. La nature de ces contraintes diffère radicalement des précédentes. D'une part, nous pouvons désormais nous abstraire de la linéarité du texte : contrairement à une approche par expressions régulières, nous pouvons ici ignorer un certain nombre d'éléments du flot textuel. D'autre part, s'il existe bien des contraintes interprétatives entre l'introducteur et certains éléments de la zone introduite, ces contraintes n'imposent pas un ordre strict entre ces éléments. Pour l'expression de telles contraintes à la fois *non linéaires* et *non séquentielles*, nous disposons du formalisme CDML et pouvons formuler la « grammaire » ci-dessous :

```
Unit cadre {type:"cadre"}:
  start(pattern: {type:"introduceur"})
  homogeneity(comparator: portée)
Comparator portée ({type:"verbe"} as $v1,
  {type:"verbe"} as $v2):
  $v1/temps = $v2/temps
Comparator portée ({type:"introduceur"} as $i,
  {type:"tempo"} as $t):
  (($i/debut >= $t/debut) and ($i/debut <= $t/fin))
or
  (($i/fin >= $t/debut) and ($i/fin <= $t/fin))
```

Nous recherchons une unité textuelle commençant par un élément identifié comme « introducteur » dont tous les verbes sont au même temps, et au sein de laquelle les expressions temporelles portent sur un intervalle compatible avec l'introducteur.

La figure 4 illustre les différentes étapes de l'analyse des cadres temporels, en précisant, pour chacune d'elles, le « grain » auquel elle opère et le modèle d'analyse utilisé. Pour une présentation plus détaillée de cet analyseur de cadres temporels et de son évaluation, on pourra se reporter à Ferrari *et al.*, 2005.

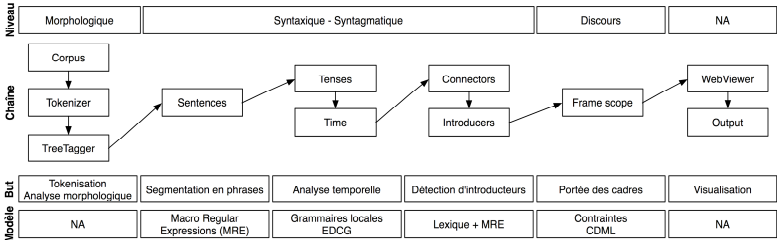


Figure 4 - La chaîne d'analyse des cadres temporels de discours

Il est ainsi possible, à l'aide des principes méthodologiques promus par la plate-forme, et en nous appuyant sur la complémentarité des modèles d'analyse, de mettre en place un analyseur de cadres temporels, certes encore imparfait, mais ne faisant usage que de formalismes purement déclaratifs propices à la capitalisation de l'expertise linguistique.

5. Perspectives

Pour l'heure, la plate-forme s'adresse principalement à des informaticiens linguistes et des linguistes déjà familiarisés avec les approches formelles, pour une utilisation optimale des modèles d'analyse proposés. La constitution déjà entreprise d'une communauté d'utilisateurs d'horizons variés, et un effort significatif de documentation devra permettre de pallier progressivement cette restriction. D'autre part, la mise à disposition de ressources (lexiques, grammaires...) dédiées à différentes tâches fait actuellement l'objet d'un important travail. Enfin, l'intégration de nouveaux outils et de nouveaux modèles d'analyse permettra d'aborder de nouveaux problèmes (fouille de texte, cartographie de corpus...) et d'optimiser certains traitements déjà proposés (analyse morphologique et syntaxique en particulier).

Références

Références bibliographiques

- BILHAUT, Frédéric & WIDLÖCHER, Antoine (2006a). LinguaStream : An Integrated Environment for Computational Linguistics Experimentation, in *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL) (Companion Volume)*, Trento, Italy, 95-98.
- BILHAUT, Frédéric & WIDLÖCHER, Antoine (2005a). La plate-forme LinguaStream, Journée ATALA « Articuler les traitements sur corpus », Paris, France (Working Paper).
- BILHAUT, Frédéric, HO DAC, Lydia-Mai, BORILLO, Andrée, CHARNOIS, Thierry, ENJALBERT, Patrice, LE DRAOULEC, Anne, MATHET, Yann, MIGUET, Hélène, PÉRY-WOODLEY, Marie-Paule & SARDA, Laure (2003). Indexation discursive pour la navigation intradocumentaire : cadres temporels et spatiaux dans l'information géographique, in *Actes de la 10e Conférence Traitement Automatique du Langage Naturel (TALN)*, Batz-sur-Mer, 315-320.
- BOURIGAULT, Didier & FABRE, Cécile (2000). Approche linguistique pour l'analyse syntaxique de corpus, *Cahiers de grammaire*, 25 : 131-151.
- CHAROLLES, Michel (1997). L'encadrement du discours : Univers, champs, domaines et espaces. *Cahiers de Recherche Linguistique*, 6. Université de Nancy 2.
- CUNNINGHAM, Hamish, MAYNARD, Diana, BONTCHEVA, Kalina & TABLAN, Valentin (2002). GATE: A framework and graphical development environment for robust NLP tools and applications, in *Proceedings of the 40th Meeting of the Association for Computational Linguistics*.
- FERRARI, Stéphane, BILHAUT, Frédéric, WIDLÖCHER, Antoine & LAIGNELET, Marion (2005). Une plate-forme logicielle et une démarche pour la validation de ressources linguistiques sur corpus : application à l'évaluation de la détection automatique de cadres temporels, in *Actes des 4èmes Journées de Linguistique de Corpus*, Lorient, France (à paraître).
- MULLER, Claude, ROYAUTE, Jean & SILBERZTEIN, Max (2004). *INTEX pour la Linguistique et le Traitement Automatique des Langues*. Presses Universitaires de Franche-Comté.
- SCHMID, Helmut (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees, in *Proceedings of the Conference on New Methods in Language Processing*, Manchester, UK.
- WIDLÖCHER, Antoine (2006). Analyse par contraintes de l'organisation du discours, in *Actes de TALN 2006*, 367-376, Leuven, Belgique.
- WIDLÖCHER, Antoine & BILHAUT, Frédéric (2005). La plate-forme LinguaStream : un outil d'exploration linguistique sur corpus, in *Actes de la 12e Conférence Traitement Automatique du Langage Naturel (TALN)*, Dourdan, 517-522.

Références aux sites Internet et logiciels

- BILHAUT, Frédéric & WIDLÖCHER, Antoine (2006b). The LinguaStream Platform website, <http://www.linguastream.org>, septembre 2006.
- BILHAUT, Frédéric & WIDLÖCHER, Antoine (2005b). *LinguaStream 2.0*. Caen: GREYC.

Autour des langues et du langage est un ouvrage réunissant une sélection d'articles issus du Colloque international des Étudiants-chercheurs en Didactique des Langues et en Linguistique, qui a eu lieu à Grenoble en juillet 2006. Il présente les recherches actuelles de jeunes chercheurs du monde entier et reflète la diversité des objets, des méthodes et des perspectives dans l'étude des langues et du langage.

Les contributions ont été regroupées en 5 axes thématiques : Description linguistique, Psycholinguistique et développement langagier, Sociolinguistique et plurilinguisme, Didactique des langues et Outils informatiques pour la linguistique et la didactique des langues. Au delà de ce choix éditorial, c'est la pluridisciplinarité qui est mise à l'honneur dans ce recueil, destiné aux étudiants, aux enseignants-chercheurs et à tout lecteur curieux d'élargir ses connaissances à travers ce panorama riche et varié.